



University
of Glasgow

<https://theses.gla.ac.uk/>

Theses Digitisation:

<https://www.gla.ac.uk/myglasgow/research/enlighten/theses/digitisation/>

This is a digitised version of the original print thesis.

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study,
without prior permission or charge

This work cannot be reproduced or quoted extensively from without first
obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any
format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author,
title, awarding institution and date of the thesis must be given

Enlighten: Theses

<https://theses.gla.ac.uk/>
research-enlighten@glasgow.ac.uk

A THEORETICAL STUDY OF QUANTUM BALLISTIC TRANSPORT IN SEMICONDUCTOR RING STRUCTURES.

**By
Michael Finch.**

A thesis
presented to the University of Glasgow in partial fulfilment of
the requirements for entry to the degree of Doctor of Philosophy.
Department of Electronics and Electrical Engineering.
March 1989.

© M. Finch 1989.

ProQuest Number: 10999341

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10999341

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Table of contents.

Table of contents	(i)
List of Illustrations	(vi)
Acknowledgements	(x)
Abstract	(xi)
Introduction.	1
Motivation and purpose	1
Synopsis of thesis	3
Chapter 1. Quantum interference phenomena.	7
1.1 Introduction.	7
1.2 Classes of quantum devices.	7
1.3 The Aharonov-Bohm effect.	9
1.3.1 General.	9
1.3.2 The Aharonov-Bohm effect in vacuo.	12
1.3.3 The Aharonov-Bohm effect in metals.	13
1.3.4 The Aharonov-Bohm in semiconducting structures.	16
1.4 The characteristics of different types of quantum interference processes.	17
1.5 Summary.	21
Chapter 2. Quantum-mechanical modelling of devices in one dimension.	22
2.1 Introduction.	22
2.2 Methods of calculating the transmission coefficient of a one-dimensional potential.	22
2.3 Transmissi matrix methods.	24
2.4 The use of the transmission matrix to study a resonant tunnelling device.	26
2.4.1 Traversal times.	30
2.4.2 Results of tunnelling delay time calculations for a resonant double-barrier potential.	31
2.5 A group-theoretical approach to the transmission matrix method.	32
2.6 Summary.	37

Chapter 3.	Theoretical origin of the AB effect.	3 8
3.1	Introduction.	3 8
3.2	Gauge invariance.	3 9
3.3	A simple analysis of the AB effect.	4 3
3.4	The electric AB effect.	4 6
3.5	Berry's geometric phase factor.	4 8
3.6	Summary.	5 0
 Chapter 4.	 Structures achieving confinement of carriers in one and two dimensions.	 5 1
4.1	Introduction.	5 1
4.2	Transport phenomena in the inversion layers of silicon NMOS transistor structures.	5 1
4.2.1	General.	5 1
4.2.2	Transport in a multi-terminal MOSFET.	5 2
4.2.3	Electrostatic confinement of the 2DEG in an NMOS structure.	5 3
4.3	GaAs HEMT structures.	5 4
4.3.1	General.	5 4
4.3.2	Operational principles.	5 4
4.3.3	Confinement of the 2DEG in HEMT structures	5 6
4.4	Summary.	5 8
 Chapter 5.	 Modelling of an AB ring with one-dimensional conductors.	 5 9
5.1	Introduction.	5 9
5.2	Elements of the model.	5 9
5.3	The transmission coefficient of a ring with one-dimensional conductors.	6 2
5.4	Results.	6 8
5.5	Discussion of results: a phasor diagram approach.	6 9
5.6	A qualitative study of the effects of impurities in one dimension.	7 2
5.7	Summary.	7 4

Chapter 6. Two-dimensional time-dependent transport. 75

6.1	Introduction.	75
6.2	Motivation for a 2D model.	76
6.3	Specification of the problem to be solved.	76
6.4	Solution of the time-dependent Schrodinger equation.	78
6.5	The stability of expansion schemes for the time-evolution operator.	80
6.6	The importance of unitarity in the expansion scheme.	81
6.7	Numerical solution of the equations.	83
6.8	Solution of the tri-diagonal matrix equations.	84
6.9	Generalisation of the solution technique for two dimensions.	86
6.10	Errors arising from the non-commutivity of the split-operators.	87
6.11	Vector-processing techniques employed for high-speed computation.	92
6.11.1	Modification of computer code for vector processing.	93
6.12	Summary.	97

Chapter 7. Computer modelling of an AB ring. 98

7.1	Introduction.	98
7.2	Construction of the Hamiltonian.	98
7.3	Construction of the Hamiltonian for the case of a magnetic string piercing the centre of an AB ring.	100
7.4	Specification of the initial condition.	102
7.5	Errors incurred in the process of discretisation.	103
7.6	Choice of an appropriate timestep.	106
7.7	Errors in the phase of the wavefunction.	107
7.8	Modelling the contacts.	107
7.9	Summary.	110

Chapter 8.	Results of the two-dimensional simulations.	111
8.1	Introduction.	111
8.2	Description of the idealised AB ring modelled	111
8.3	Preliminary results.	112
8.4	A full-scale computer model.	114
8.4.1	General analysis of results for a large ring	117
8.4.2	Multimode structure.	118
8.4.3	Transit times.	122
8.5	Summary.	125
Chapter 9.	Magneto-resistance studies.	126
9.1	Introduction.	126
9.2	Determination of transmission and reflection coefficients.	127
9.3	The relation between the transmission coefficient and the conductance.	127
9.4	The resistance of small rings.	130
9.5	The magneto-transmission characteristic of a 400nm diameter ring.	131
9.6	Comparison with one-dimensional analysis and experimental results.	133
9.6.1	General.	133
9.6.2	Analysis of magneto-transmission results in terms of hybrid 1D and double-slit models.	134
9.6.3	Harmonic analysis of simulated magneto-transmission characteristics.	136
9.6.4	Comparison of simulated results with experiment.	138
9.7	A method of increasing the modulation depth of the resistance oscillations.	141
9.8	Other factors influencing the magneto-transmission characterisitic of an AB ring.	143
9.8.1	The effect of the fluctuation potential caused by remote donors.	143
9.8.2	The effect of a longitudinal electric field.	145
9.9	A fluid dynamical analogy.	146
9.9.1	General.	146
9.9.2	Effect of an imaginary potential in the fluid model.	147
9.9.3	Consequences of the quantum potential in narrow conductors.	148
9.10	Summary.	150

Conclusion of thesis.	152
Appendix A.	The transmission matrix. 159
	Definition. 159
	Properties of the T-matrix. 161
	Relation to an equivalent scattering matrix. 162
	Calculation of matrix elements for a rectangular potential barrier. 163
Appendix B.	The stationary phase approximation. 166
Appendix C.	Estimation of wavepacket energy from the sheet carrier concentration. 168
References.	172

List of illustrations.

<u>Figure number.</u>		<u>Following page.</u>
1.1	Classes of quantum devices.	9
1.2	Schematic diagram of wavepacket propagation around a ring.	On page 10
1.3-1.5	Double-slit interference experiments	11
1.6	Biprism of Mollenstedt and Duker.	12
1.7	Diagram of iron whisker used in interference experiments.	12
1.8	Qualitative reproduction of Chambers' results.	12
1.9	Sharvin and Sharvin's Mg cylinder.	14
1.10	$h/2e$ oscillations obtained by Sharvin and Sharvin.	14
1.11	Loop fabricated by Umbach <i>et al</i> to search for h/e oscillations.	16
1.12	Results from the loop of Umbach <i>et al</i> .	16
1.13	Clear observation of h/e oscillations by Webb <i>et al</i> .	16
1.14	Proposed semiconducting structure making use of the electrostatic AB effect.	16
1.15	The different types of interference processes and how they are related.	20
1.16	Magneto-resistance data obtained by Leadbeater <i>et al</i> from semiconductor wires.	20
1.17	Fourier transforms of the data in figure 1.16	20
2.1	General T-matrix problem.	25
2.2	Type of potential which is well-suited for T-matrix methods.	25
2.3	Model potential used for the calculation of T-matrix elements.	25
2.4	Transmission versus energy plot for a double barrier potential.	29
2.5	Tunnel delay time versus energy plot for a double barrier potential.	29
3.1	The generic AB experiment.	45
3.2	The principle of the electrostatic AB effect.	45
3.3	Diagram of an AB ring under the influence of a transverse electric field.	47
3.4	Parallel transport of a vector over a curved surface.	47
4.1	Diagram of Skocpol's multi-terminal MOSFET.	53

4.2	Resistance of a channel segment of Skocpol's device.	53
4.3	Creation of a Q1D channel by electrostatic confinement.	53
4.4	Fluctuation phenomena in a Q1D channel.	53
4.5	Conduction band profile of a HEMT.	54
4.6-4.9	Techniques used to provide lateral confinement of the 2DEG.	57
4.10	An early AB ring fabricated by Ford <i>et al.</i>	57
4.11	Cross-sectional view of 4.10	57
4.12	An Improved ring structure fabricated by Ford <i>et al.</i>	57
5.1	One-dimensional modelling of an AB ring.	62
5.2-5.5	Diagrams showing the matching of the plane-wave coefficients using S and T matrices. On pages 63,64,65	
5.6-5.9	Transmission versus kl product for a one-dimensional ring (fixed coupling parameter, variable flux).	73
5.10 & 5.11	Transmission versus kl product for a one-dimensional ring (fixed flux, variable coupling parameter)	73
5.12-5.14	Transmission versus flux for a one-dimensional ring (fixed coupling parameter, variable kl product).	73
5.15 & 5.16	Transmission versus flux for a one-dimensional ring (fixed kl product, variable coupling parameter)	73
5.17 & 5.18	Diagrams showing the interpretation of the transmission plots in terms of phasors.	73
5.19 -5.21	Effect of elastic scattering in one arm of the ring.	73
5.22-5.24	Effect of elastic scattering in both arms of the ring.	73
5.25-5.27	Dependence of the transmission coefficient on the coupling parameter in the presence of scattering.	73
5.28 & 5.29	Two sample plots demonstrating the wide range of behaviour possible when scattering is introduced.	73
6.1	Restructuring of the computer algorithm to allow vector processing.	94
8.1	Idealised HEMT structure used in computer modelling.	113
8.2-8.4	Time evolution of a wavepacket incident on a 30nm ring (no flux applied).	113
8.5-8.7	Time evolution of a wavepacket incident on a 30nm ring (magnetic string of strength $1/2(h/e)$ applied to centre).	113
8.8 & 8.9	3D representations of the wavefunction.	113

8.10-8.12	Time evolution of a wavepacket incident on a 30nm ring (uniform field applied).	1 1 3
8.13-8.33	Full-scale simulation of a wavepacket incident on a 400nm ring (no field).	1 1 7
8.34	Complex form of the wavefunction distribution when the ring is modelled as a closed system.	1 1 7
8.35	Chaotic behaviour obtained by MacDonald and Kaufman	1 1 7
8.36-8.55	Simulation of a wavepacket incident on a 400nm ring (half a flux quantum applied through ring)	1 1 7
8.56-8.60	3D representations of the wavefunction as the wavefront reaches the output.	1 1 7
8.61-8.80	Simulation of a wavepacket incident on a 400nm ring with an absorber in one arm (no field).	1 2 1
8.81	Wavefront position versus time for a 10meV wavepacket propagating around a 400nm ring.	1 2 4
8.82	Wavefront position versus time for a 1.27meV wavepacket propagating around a 400nm ring.	1 2 4
8.83	Transmitted charge position versus time for a 10meV wavepacket propagating around a 400nm ring.	1 2 4
8.84	Transmitted charge position versus time for a 1.27meV wavepacket propagating around a 400nm ring.	1 2 4
9.1	Calculation of the conductance of a 1D region using the Landauer formula.	1 3 2
9.2	Magneto-transmission characteristic of a 400nm ring for a 10meV wavepacket energy.	1 3 2
9.3	Simulated magneto-transmission characteristic for the ring of Ford <i>et al</i> using a wavepacket energy of 9meV	1 3 5
9.4	Simulated magneto-transmission characteristic for the ring of Ford <i>et al</i> using a wavepacket energy of 13meV	1 3 5
9.5	1D transmission characteristic using model parameters obtained from the 2D simulations.	1 3 5
9.6-9.11	Hybrid 1D and double-slit models for an AB ring.	1 3 5
9.12	Fourier transform of one cycle of the magneto-transmission curve for the ring of Ford <i>et al</i> using a 9meV wavepacket.	1 4 0

9.13	Fourier transform of one cycle of the magneto-transmission curve for the 400nm ring using a 10meV wavepacket.	140
9.14	Fourier transform of two cycles of the magneto-transmission curve for the ring of Ford <i>et al</i> using a 9meV wavepacket.	140
9.15	Fourier transform of two cycles of the magneto-transmission curve for the 400nm ring using a 10meV wavepacket.	140
9.16	Sample of experimental data obtained by Ford <i>et al</i> .	140
9.17	Magneto-resistance characteristic of the ring of Ford <i>et al</i> using a 9meV wavepacket.	140
9.18	Magneto-resistance characteristic of the ring of Ford <i>et al</i> using a 13meV wavepacket.	140
9.19	Magneto-transmission curve obtained for a ring with conductors of 200nm width.	140
9.20	Magneto-resistance curve obtained for a ring with conductors of 200nm width.	140
9.21-9.40	The effect of constricting the exit of the ring (no applied flux)	142
9.41-9.46	The effect of constricting the exit of the ring (half a flux quantum applied)	142
9.47	Typical magneto-transmission characteristic in the presence of a random fluctuation potential.	144
9.48	Effect of a longitudinal electric field on the on/off transmission ratio.	145
A1	Calculation of a transmissison coefficient using T-matrices.	165
A2	Matching the plane-wave coefficients for a model1D potential	165
C1-C3	Density of states versus energy for conductors of 20, 66.4 and 200nm width respectively.	171
C4-C6	Number density versus Fermi energy for conductors of 20, 66.4 and 200nm width respectively.	171
Table 7.1	Gaussian truncation errors.	104

Acknowledgements.

There are many people who have helped me both directly and indirectly with this project over the past three years whom I would like to thank.

My Parents for their encouragement and constant support.

Professor J.R Barker for many stimulating discussions and for his enthusiasm in supervising this project.

All those in the theory group: Dr. M. Al-Mudares, Dr. J.H. Davies, Dr. J. Nixon, M. Laughton, and especially J. Pepin for their willingness to discuss my technical problems and for their friendship.

The computing service staff in both the Engineering Department and the Central Computing Service, of whom I would particularly like to thank: D.D Campbell, A.I MacKinnon and Dr. P. Rosenberg.

Thanks also to Dr. J. Hague of IBM (UK) for his help on vector-processing techniques.

I am also most grateful to C.J.B Ford for supplying me with some of his experimental results and for many helpful discussions.

I would also like to thank C.J.S. Bowden, J.P. Gibson, P. Anwyl, M. Binet, J. Binet, R. Chaudhry, M. McCarl, A.S. Duffy, S. Gallagher, L. Gallagher.

Abstract.

Recent developments in microfabrication technology have enabled the manufacture of semiconductor devices in which the carriers scatter very infrequently over typical device lengths. Transport of this kind is termed ballistic, and under such conditions, coherent quantum interference phenomena become an increasingly important part of the conduction process. In particular, the conductors of such devices now assume the role of electron waveguides.

Most previous attempts at modelling quantum ballistic transport have been based on one-dimensional models. However, relatively little was known about the true nature of wavepacket propagation in real structures where diffraction from apertures or around obstacles could occur.

This thesis presents the first theoretical study of quantum ballistic transport in a two-dimensional quantum waveguide network. The study specifically concentrates on modelling the Aharonov-Bohm effect in ring structures, which is an exclusively quantum-mechanical effect.

The method of investigation was to numerically solve the two-dimensional time-dependent Schrodinger equation for an idealised ring structure using a computer algorithm which incorporated several novel techniques.

One-dimensional calculations show that one can expect a modulation depth of 100% in the oscillations in the magneto-resistance characteristic of such rings. Present oscillation amplitudes measured experimentally however fall far short of this figure, typically being about 0.1% of the background resistance in metal rings and about 10% in rings formed in the two-dimensional electron gas at a heterojunction interface.

Computer simulation of wavepacket propagation in these latter structures clearly show a multi-mode structure in the wavefunction across the conductors of realistically-sized rings. It is shown that it is the transmission of more than one mode at the exit of the ring which is a major factor in reducing the amplitude of the magneto-resistance oscillations. Good agreement between the average magneto-resistance oscillation amplitude in the simulated and experimental characteristics for a ring formed at a heterojunction was obtained. The two-dimensional model can therefore be regarded as a major improvement on earlier one-dimensional models.

Evidence suggesting a damping of the magneto-resistance oscillations as a result of the direct action of the magnetic field acting on the conductors is also found. It is estimated that the approximate cut-off field would be about 0.5 Tesla for the particular device modelled, which is consistent with experimental observations of a decline in the oscillation amplitude in the range 0.5-1.0 Tesla.

A modification of the basic ring structure to achieve larger magneto-resistance oscillations by constricting the exit of the ring is proposed and computer simulation of wavepacket propagation through this structure shows that a substantial increase in modulation depth can be expected.

The techniques developed in this thesis have therefore been able to successfully model existing quantum interference devices and also assess the likely improvement in performance of a hypothetical device. These techniques could also be applied to the modelling of wavepacket propagation in other types of sub-micron quantum-interference devices where transport can be considered to be ballistic.

INTRODUCTION.

Motivation and Purpose.

Semiconductor physics and technology is now entering an exciting era in which new phenomena are being discovered and harnessed to achieve novel types of electrical and optical devices.

Much of this activity has been based on the modifications to the bulk material transport properties which can be brought about by artificially tailoring the bulk bandstructure at a microscopic level. Other novel effects have arisen from continued miniaturisation of electronic components where quantum-mechanical behaviour is becoming a significant aspect in the overall conduction process.

Although many such phenomena have been known to theoretical physics for some time, the past decade has seen an intensification of research in these areas. This is because the types of structures required for these effects to become prominent can now be realised following constant improvements in microfabrication technology. These improvements have fuelled theoretical research in this area which, in turn, has created the feedback to stimulate further advances in fabrication technology.

If microelectronics is to enjoy the further benefits of miniaturisation, the fluctuation phenomena caused by the non self-averaging of random processes at a microscopic level must be properly understood. However, a more optimistic view of these processes is that they can be used to probe the underlying physics of the conduction process at a microscopic scale. Indeed, quantum-mechanical interference can form the basis of new types of devices where the small size is an advantage.

The fluctuation phenomena peculiar to conducting rings such as the Aharonov-Bohm (AB) effect and the closely-related Al'tshuler, Aronov and Spivak (AAS) effect are due to interfering electron trajectories following different paths which are explicitly defined by the conductors. Interference phenomena such as weak-localisation and universal conductance fluctuations on the other hand can occur in single wires. These processes are also due to interfering electron trajectories following different paths in a two-dimensional plane but within the wire itself.

Theoretical understanding of such phenomena in the collision-dominated diffusive regime, which occurs mainly in metals and highly-doped semiconductors, is well advanced. There is now however increasing attention being focussed on these interference effects, particularly the AB effect, in channels within the high-mobility 2DEG formed at a heterojunction. In this case electrons scatter infrequently over typical device lengths and transport within such devices can be considered to be ballistic with the conductors now assuming the role of electron waveguides. Relatively little was known theoretically about this type of transport as it occurred in real ring structures.

Many early attempts to model this behaviour were one-dimensional and based on the matching of plane wave coefficients around the ring. It soon became apparent however that these models had the serious shortcoming that they predicted a 100% modulation in the resistance of the ring with applied magnetic field when every experimental result to date was much less than this, typically being about 10% of the background value at most.

Real rings do not have one-dimensional conductors and the electron wavepacket has some freedom to diffract out of apertures and around structures in two (or three) dimensions. It was therefore believed that a more realistic study of ballistic transport in these structures could be achieved by considering a two-dimensional model of wavepacket propagation in these structures. In particular it was intended to discover why the magnitude of the oscillations were so small and what effect the ring geometry would have on the magneto-resistance oscillation amplitude.

Synopsis of thesis.

The first chapter describes how quantum interference processes can affect the bulk transport properties of semiconductors. It is shown how different types of devices can be classified in terms of their dimensionality arising from varying degrees of carrier confinement. The process of major concern in this thesis is the Aharonov-Bohm effect in multiply-connected geometries, and the chapter contains a review of various experiments which demonstrate the effect in vacuo, metals and semiconductors together with the relevant theory.

A parameter of particular interest is the quantum-mechanical transmission coefficient of a particular structure, as this determines the current flow through the device. Chapter two considers a variety of ways this can be evaluated in one dimension but concentrates on two methods. One using transmission matrices and a more sophisticated method using group-theoretical techniques. It is shown how the transmission matrix method can be readily applied to calculate both the transmission coefficient and, via the stationary-phase approximation, the transit time for electrons incident on rectangular-shaped potentials in one dimension. The results obtained are compared with well-known results for a resonant double-barrier potential.

Chapter three concentrates on the AB effect and explains how it arises as a consequence of gauge invariance. In order to maintain gauge invariance of the Schrodinger equation the phase of the wavefunction changes by an amount determined by the four-dimensional electromagnetic flux enclosed by the electron-beam trajectories. The derived relations are used in a simple model to determine the resulting wavefunction when the two beams are allowed to interfere. This is done for both the magnetic and electrostatic cases, although they both have a common origin in the general theory.

Much recent research on the AB effect and other quantum interference phenomena has been concerned with the transport in the two-dimensional electron gas (2DEG) present in the inversion layers of silicon MOSFET devices and at heterojunction interfaces in GaAs high electron mobility transistors (HEMTs). The major interest in HEMT structures in relation to this thesis derives from the fact that the 2DEG can be further confined to form quasi-1D wires and ring structures. Concentrating mainly on GaAs devices, chapter four discusses several techniques for achieving this extra confinement and the advantages and disadvantages of each method are outlined.

A detailed description of a patterned-gate HEMT ring structure fabricated by Ford *et al* [1987] is given, as this was one of the structures for which computer simulations were performed.

Chapter five describes how an idealised AB ring with one-dimensional conductors was modelled. These studies formed the background work for the project and were used for the interpretation of the two-dimensional results obtained later. This time-independent model used transmission and scattering matrices to match the plane-wave coefficients around the ring to find the total transmission. The results confirmed the occurrence of the AB effect in one dimension and verified that the transmission fell to zero when the enclosed magnetic flux was equal to half the flux quantum h/e . The principal effect of constraining the electrons to move along one-dimensional conductors was to introduce an energy dependence into the expression for the transmission coefficient which was absent from the result obtained from the analysis of a simple two-slit interference experiment.

The behaviour of the ring could be broadly classified according to whether the ring was well or poorly coupled to its external leads. The results were in good agreement with similar work by Buttiker *et al* [1984] but the underlying physics giving rise to the transmission characteristic of a well-coupled ring with no scattering in the arms was unclear. To explain the behaviour, an approach using phasor diagrams was taken which showed that the transmission could reach zero at certain wavevector values even though the magnitude of the wavefunction had not been reduced to zero.

Real rings have random potentials arising from remote donor ions superimposed on the confinement potential and it was considered important to obtain an assessment of the likely effects of this. The one-dimensional model was therefore used to qualitatively study the effects of a scattering potential by considering the transmission coefficient of the arms of the ring to be less than unity. The results of this study showed that even a modest amount of scattering in the arms can have a large effect on the transmission characteristic of the ring.

Chapter six discusses in general terms the modelling of quantum ballistic transport in two dimensions and discusses some methods for the solution of the numerical problem. The numerical solution of the one-dimensional problem is discussed in detail as the solution of this later forms the basis for solutions of the problem in two dimensions.

The problem was ultimately reduced to the evaluation of several sets of tri-diagonal matrix equations for which standard numerical methods could be used. However, in order to fully utilise the high-speed vector processing capability of the IBM 3090 (which was used for the magneto-transmission calculations), it was necessary to restructure the numerical algorithm in a non-standard way. This technique is described in detail.

Having introduced the general solution technique, chapter seven then deals specifically with the modelling of a two-dimensional AB ring. The Hamiltonian operators for two different magnetic field distributions are derived and the errors incurred in the discretisation process are considered. The problem of unphysical reflections from the mesh boundaries was solved by the use of an imaginary potential to model an absorbing contact.

Chapter eight presents the results of the two-dimensional simulations. The results of a small-scale preliminary simulation are shown and it is described how this model was scaled-up to a more realistic size following an improvement in computing facilities. The results from the larger model are discussed in detail and it is shown that the formation of transverse modes in the conductors is a major factor limiting the magnitude of the magneto-resistance oscillations.

In order to quantify results and enable comparison with experiment, it was necessary to calculate the magneto-resistance characteristics of the AB rings. Chapter nine describes how this calculation was performed and analyses the results by comparison with the 1D models and by Fourier analysis of the magneto-transmission curves.

The first magneto-transmission calculations were performed on the idealised (though realistically-sized) ring so that the results from the earlier wavepacket propagation study could be used in the interpretation of the obtained magneto-transmission results.

The theory was compared against experiment by simulation of the magneto-resistance characteristics of a ring fabricated by Ford *et al* [1987]. Initial determinations of the magneto-resistance oscillation amplitude were rather higher than those found by experiment. However, later evidence and subsequent calculations show that this discrepancy can, to a large extent, be attributed to uncertainties in the width of the conducting channels.

Further calculations using revised values for the channel width were in better agreement with the observed oscillation amplitudes.

Based on the hypothesis that the transverse modes are responsible for the reduction of the magneto-resistance oscillations, a modified ring geometry is proposed. A simulation of wavepacket propagation through this structure showed that the on/off transmission ratio could be substantially improved.

Finally, the effects of a longitudinal electric field and the presence of remote ionised donors are qualitatively investigated.

1 Quantum interference phenomena

1.1 Introduction.

In this chapter several classes of quantum devices are described in terms of their dimensionalities. The quantum-mechanical effects of the reduced dimensionality are then examined and the different types of interference phenomena which occur are identified. Particular emphasis is placed on the Aharonov-Bohm (AB) effect which is the major concern of this thesis.

A review of experiments and accompanying theory for the AB effect in *vacuo*, metals and semiconductors is given. The AB effect in semiconductors is of particular relevance to this thesis and will be discussed in greater detail in further chapters.

1.2 Classes of quantum devices.

It may be argued that all semiconductor devices are quantum devices in the sense that many concepts of solid-state theory such as allowed energy bands, Fermi-energies and effective mass are based on quantum theory. However, usually when one refers to quantum devices, it is meant that the bulk material has been modified in such a way so that the presence of quantum effects radically alters the conventional transport properties.

Modifications to the bulk transport properties have been achieved by the use of molecular beam epitaxy (MBE), metal organic vapour deposition (MOCVD) and improvements in electron-beam lithography. Much of the current work is based on GaAs/AlGaAs materials, although there is also considerable amount of research on the InGaAs system and to a lesser extent in strained-layer silicon structures

The GaAs/AlGaAs system is particularly suitable for the manufacture of quantum devices. The low effective mass in GaAs and its composites results in a longer DeBroglie wavelength for a given electron energy, and since quantum effects become prominent only when the DeBroglie wavelength is comparable with device dimensions, this permits less stringent limits on fabrication technology.

Secondly, aluminium has a similar lattice parameter to gallium, so alternate layers of GaAs and AlGaAs with varying amounts of aluminium content can be grown whilst maintaining interfaces of good integrity. The effect of the aluminium is to increase the band-gap energy thus causing an energy discontinuity in the conduction and valence bands at each interface.

By the growth of alternate layers of GaAs and AlGaAs, the resulting conduction band consists of a series of energy barriers, which, theoretical work has shown [Collins 1986], can be interpreted in terms of an electrostatic potential. The techniques of MBE and MOCVD have been used to fabricate resonant tunnelling devices [Capasso 1986] in which the conduction band profile consists of two potential barriers of about 20-50Å thickness separated by a gap of about 20-50Å. It may be shown (chapter 2) that even if the tunnelling transmission through any one of the barriers individually is negligible, there exists a resonant energy for the double-barrier combination for which the transmission is unity. This behaviour is important because the resonant transmission can lead to a negative differential resistance in the current-voltage characteristics which could be used as the basis of a very-high-frequency amplifier or switch [Sollner *et al* 1983].

If more potential barriers are considered, more resonant or "allowed" energy levels are found to exist, which cluster together. In the limit of a large number of barriers, the device is known as a superlattice [Capasso 1986, Esaki 1986] and the clusters of resonant energies are identified with the mini conduction bands of the superlattice. Since the superlattice period is larger than the interatomic spacing of the GaAs composite, the energy bands of the superlattice are narrower than the bulk conduction bands and may consist of a series of closely-spaced energy levels; contrasting with the essentially continuous bulk conduction band.

The resonant tunnelling device and superlattice are examples in which the transport through the material has been affected by the introduction of potential barriers into the bulk conduction band which are comparable with the DeBroglie wavelength. However the definition of quantum devices can be expanded to include effects which play a more passive role in altering the conduction processes. An example of this type of device is the high electron mobility transistor (HEMT). The HEMT consists of a highly-doped AlGaAs layer grown on top of an unintentionally doped GaAs substrate.

The electrons from the the high-band-gap AlGaAs composite diffuse into the GaAs and become trapped at the interface where they form a two-dimensional electron gas (2DEG). Conventional transistor operation benefits from the higher mobility of the 2DEG which arises mainly as a result of the segregation of the conduction electrons from their donors. Thus in this case, the quantum effects are being used to enhance the performance of a conventional type of device rather than the effects themselves forming the basis of operation.

An important development of the HEMT-type structure has been the the use of shaped gate contacts which, by the formation of depletion regions, provide a means of confining the 2DEG into narrow channels and other configurations. A more complete description of these structures will be given in chapter4.

One way of classifying the various types of quantum device is by reference to their dimensionality and figure 1.1 shows how the devices discussed may be classified in this way. The patterned-gate HEMT structure is seen to be particularly versatile in that the 2DEG can be further confined to form a superlattice or resonant tunnelling device, or by the use of high confinement potentials, quasi-1D wires and Aharonov-Bohm rings. In principle the technique could also be used to fabricate arrays of loops or quantum dots.

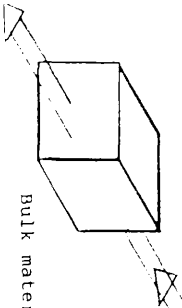
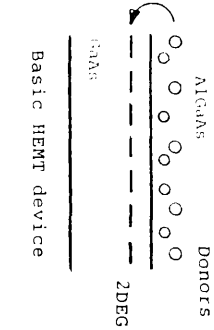
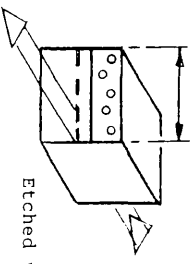
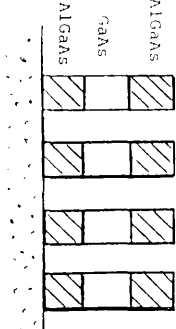
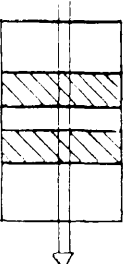
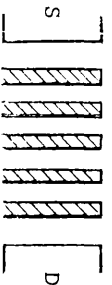
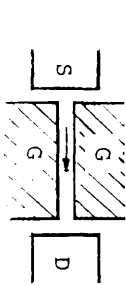
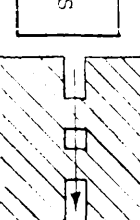
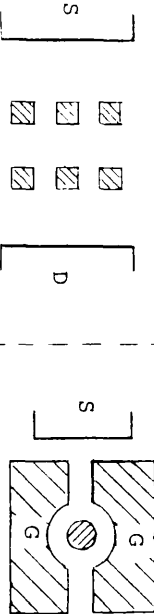
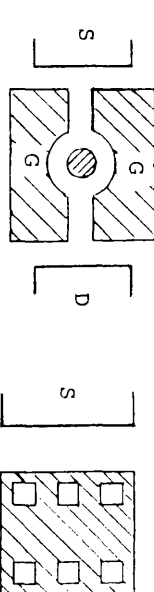
1.3 The Aharonov-Bohm effect.

1.3.1 General

Of increasing interest in the study of quantum interference effects is the quantum transport of wavepackets through ring structures and arrays of loops such as those shown in figure 1.1 [Washburn and Webb 1986].

In the case of the single ring, an incident wavepacket can be split into two coherent wavepackets propagating around each half of the ring and recombining at the opposite side where they interfere with each other.

Figure 1.1 Classes of quantum devices

	No confinement	Confinement in 1D	Confinement in 2D	Confinement in 3D
No perturbation	 <p>Bulk material</p>	 <p>Basic HEMT device</p>	 <p>Etched wires</p>	 <p>"quantum dots"</p>
+ Perturbation in 1D	 <p>RT device/Superlattice</p>	 <p>Lateral surface Superlattice</p>	 <p>Strong perturbation = Confinement</p>	
+ Perturbation in 2D		 <p>Resonant tunnelling device</p>	 <p>Patterned-gate HEMT devices</p>	 <p>Aharonov-Bohm rings</p>

A particularly interesting situation arises if the ring is subjected to a perpendicular magnetic field. The effect of the vector potential created by the field is to change the phase of the wavefunction by an amount

$$\Delta\phi = \frac{e}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} \tag{Eqn 1.1}$$

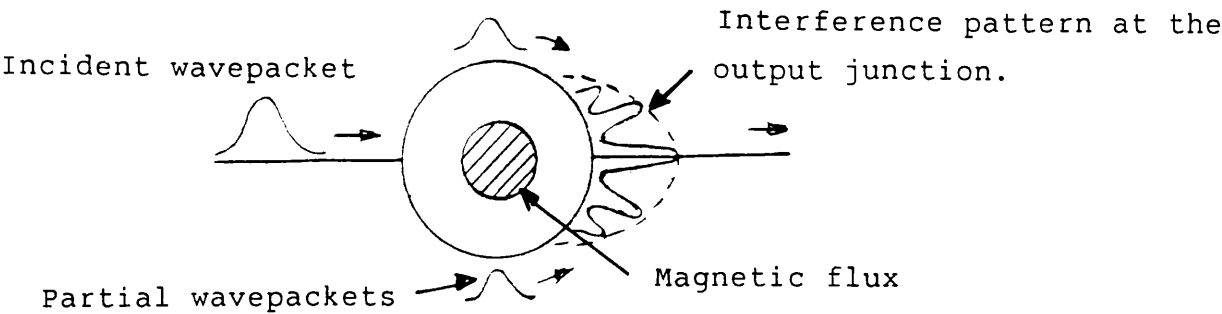
where \mathbf{A} is the vector potential.

The two partial wavepackets can therefore acquire equal but opposite additional phases due to the vector potential which means that the interference and hence current at the output can be controlled by the magnetic field. In particular, when the relative phase between the wavepackets is π , complete cancellation occurs and simple theory predicts that the output current should fall to zero. As the magnetic field is increased, the resistance of the ring is predicted to vary in a sinusoidal fashion as the wavepackets experience consecutive conditions of constructive then destructive interference. This phenomenon is known as the Aharonov-Bohm (AB) effect [Aharonov and Bohm 1959].

Figure 1.2 shows schematically how the effect occurs in ring structures.

The AB effect is closely analogous to Young's double-slit experiment in optics, and indeed many of the early experiments on the AB effect using modified electron-beam microscopes were of the double-slit type geometry.

Figure 1.2 Schematic representation of wavepacket propagation in ring structures.



Theoretical studies (discussed in greater detail in chapter 3) have predicted the following characteristics for the AB effect.

1. The resistance oscillation is a periodic function of magnetic flux threading the ring, of period h/e .
2. The interference is controlled only by the threaded flux and not directly by the field strengths.
3. The effect of a magnetic field acting on the trajectories is to displace the interference pattern as a whole, but not directly alter the arrangement of interference fringes within the overall single-slit diffraction envelope.

Figures 1.3 to 1.5 illustrate these points for a generic double-slit experiment, although the same principles also apply to the ring geometry.

The result that the interference pattern can be changed by a magnetic field which is confined solely to the non-accessible region is a surprising one because it means that the current flow through a ring structure can be changed without any force being exerted on the electrons. It also means that the electrons are aware of the presence of the magnetic field in a region which they cannot visit. It must also be stressed that this behaviour is not due to any effect arising from the close proximity of the conduction paths to the magnetic field as the conductors can, in principle, be moved arbitrarily far away from the region of magnetic field provided that the wavetrains still remain coherent.

However, from a theoretical point of view, one cannot ignore *a priori* the fact that irrespective of how far away the conductors are from the region containing the magnetic field, an extremely small, but nevertheless finite, overlap between the wavefunction and the field-containing region occurs. The same argument can also be applied to the case where the electrons are prevented from entering the "non-accessible" region by confining potentials because of the physical impossibility of infinitely large potentials. This problem has been resolved theoretically by considering the AB effect for the case of finite confinement potentials. It was then shown [Olariu and Popescu 1985] that the AB effect still occurred in the limit of the confinement potential tending to infinity at which point the space available to the electrons was truly multiply-connected.

Figures 1.3 to 1.5 Showing the effect of an enclosed flux and the direct action of a magnetic field on a double-slit diffraction pattern.

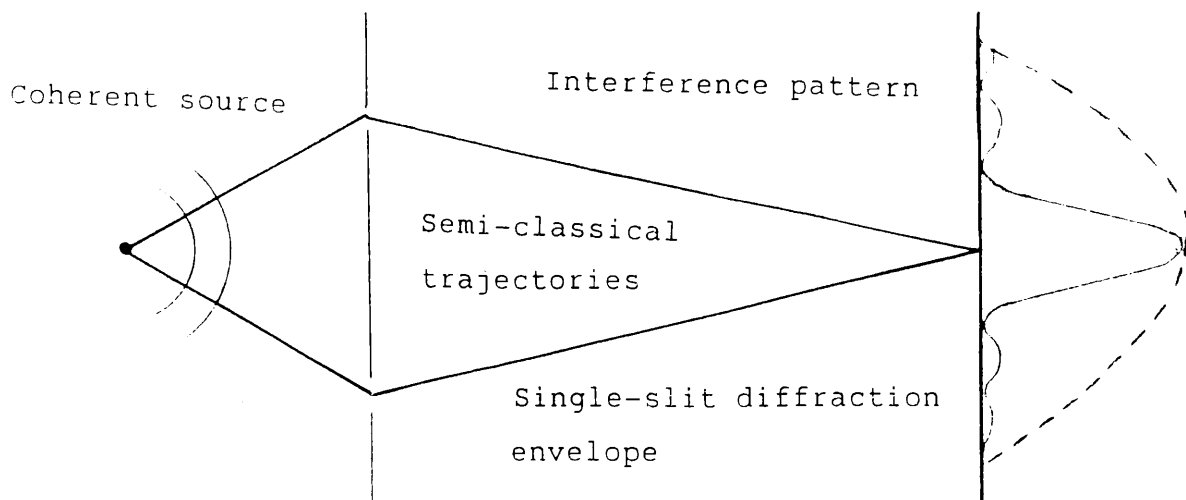


Figure 1.3 No magnetic field, no enclosed flux.

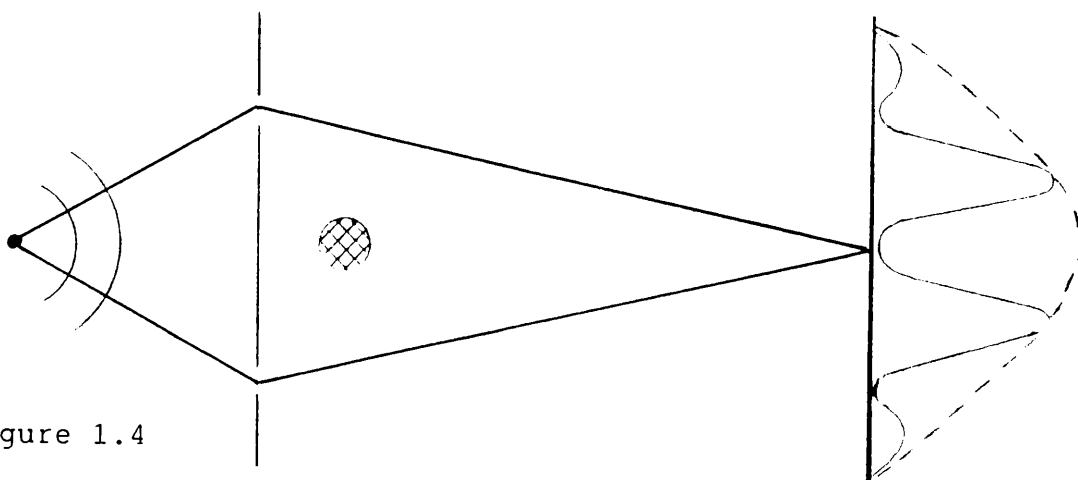


Figure 1.4

Magnetic field confined to non-accessible region.

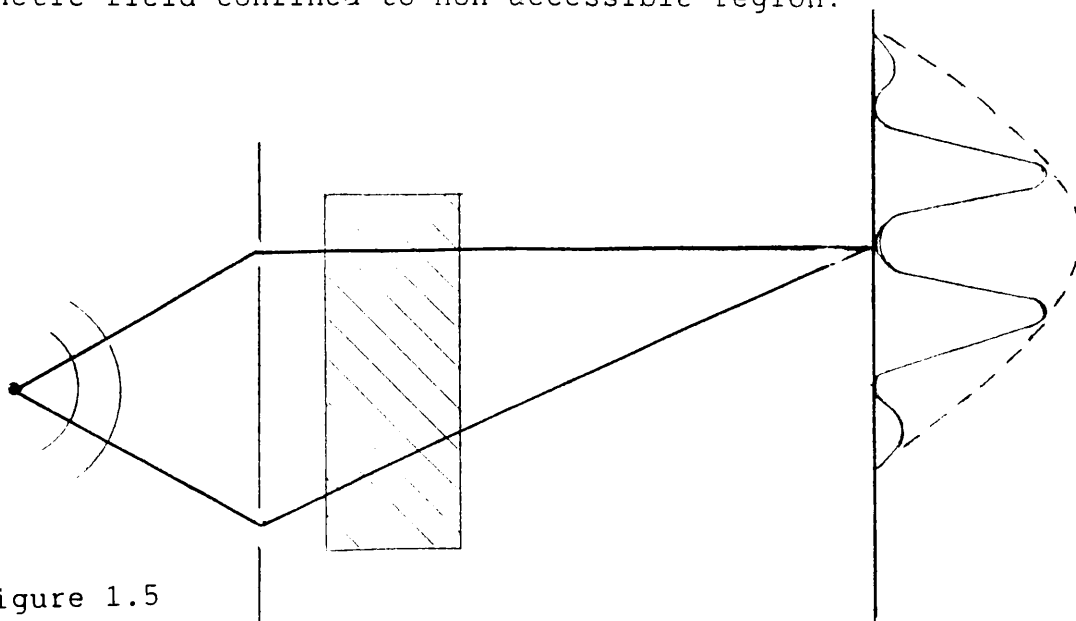


Figure 1.5

Magnetic field distribution acting on the beam trajectories and providing an enclosed flux.

In the following sections the experimental evidence for AB-type quantum interference effects in vacuum, normal-metals and semiconductors will be discussed alongside the relevant theory.

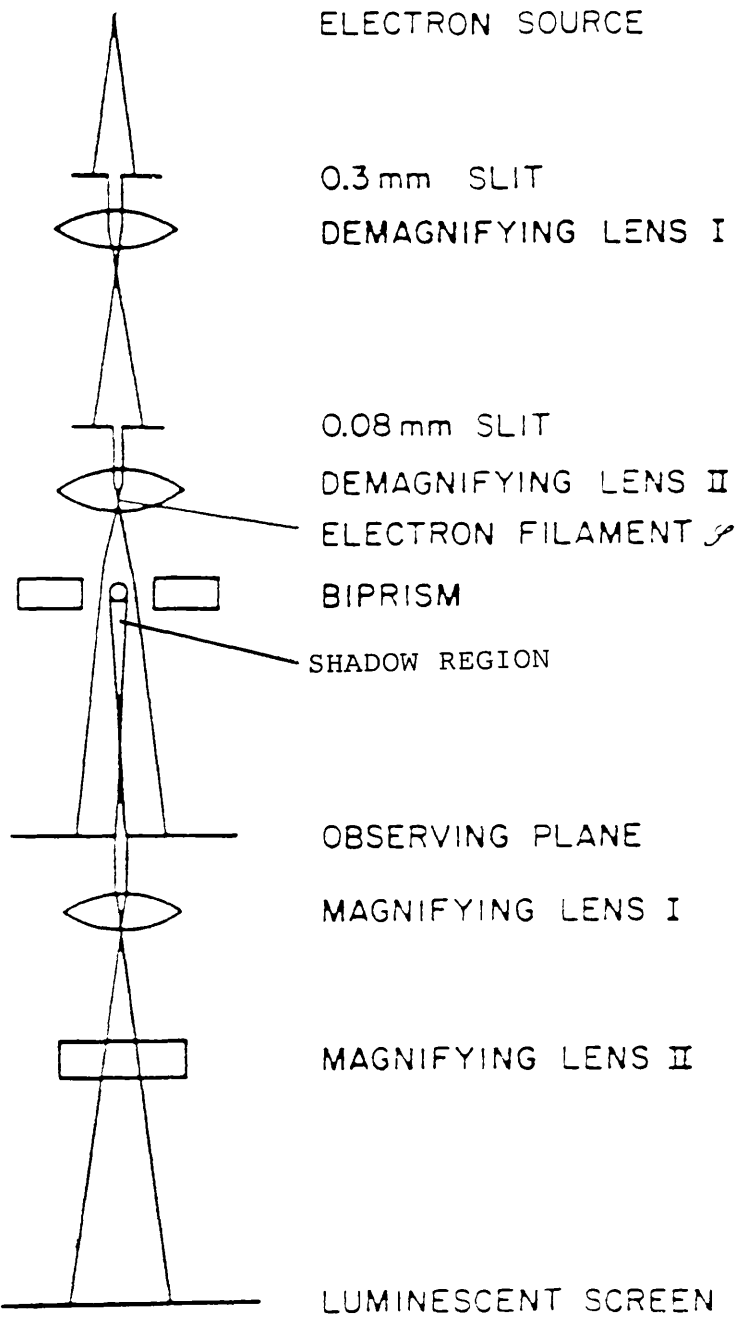
1.3.2 The Aharonov-Bohm effect in vacuo.

The quantum interference of electrons resulting from enclosed fluxes was noted by Franz [1939] as far back as 1939. Later, Ehrenberg and Siday [1949] also predicted observable quantum interference effects in the context of electron optics. Unfortunately, it seems that these two papers aroused little interest outside the immediate scope of their specialist areas. Ten years later still, Aharonov and Bohm published their widely acclaimed paper [Aharonov and Bohm 1959] on the effects of electromagnetic fluxes on quantum phenomena. The success of this paper lay in the fact that the wider implications of the role of electromagnetic fluxes in quantum mechanics was emphasized whilst also pointing out that the technology necessary for testing the theory already existed.

The earliest experiments demonstrating the AB effect used electrostatic biprisms of the sort first used by Mollenstedt and Duker [1956] to observe electron diffraction patterns around a metallic fibre. A schematic diagram of the electrostatic biprism is shown in figure 1.6 . Since contributions from different parts of the source are incoherent, it was necessary that the width of the source did not exceed about 100nm to ensure a coherent beam of electrons. This narrow source was obtained using electron-optical techniques from a real electron source of 50 microns width. The central fibre was held at a positive potential with respect to the two grounded plates and had the effect of deflecting the electron beam by a certain angle which was independent of the distance of closest approach to the fibre. The effect of this was to create two virtual sources either side of the fibre and was made possible by the fact that the $1/r$ -dependence of the electric field produced by the charged fibre balanced the r -dependence of the time spent in that electric field. Using this electron focussing geometry, Mollenstedt and Duker were able to demonstrate the quantum-mechanical analogue of Fresnel diffraction around the fibre.

Figure 1.6

Schematic diagram of the electrostatic biprism of Mollenstedt and Duker. The two demagnifying lenses are to reduce the apparent size of the electron source and the two magnifying lenses are to increase the size of the resulting interference pattern.



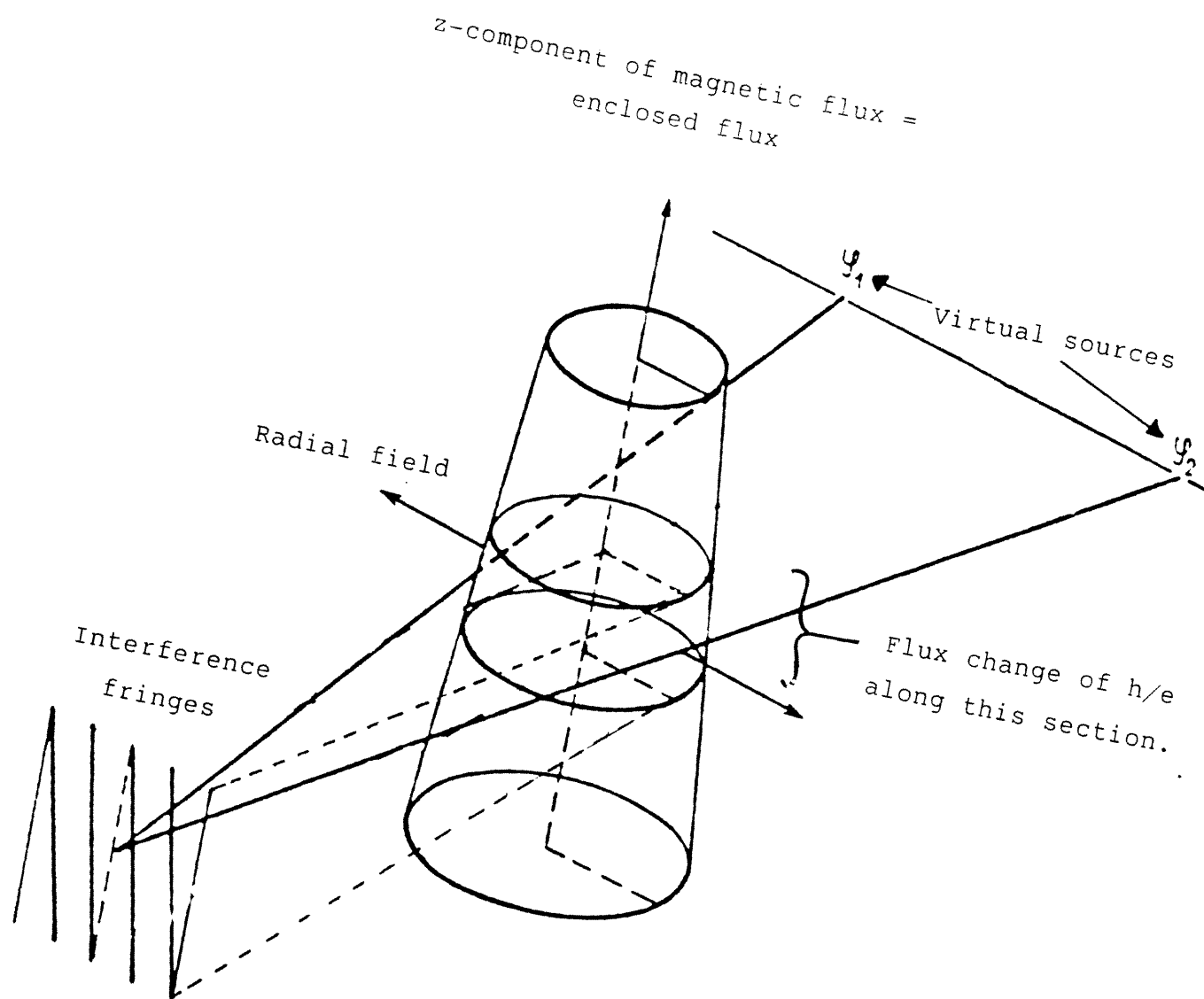


Figure 1.7

Magnetic whisker used by Chambers to demonstrate the quantum effects of an enclosed magnetic flux.

Approx. 0.6 micron

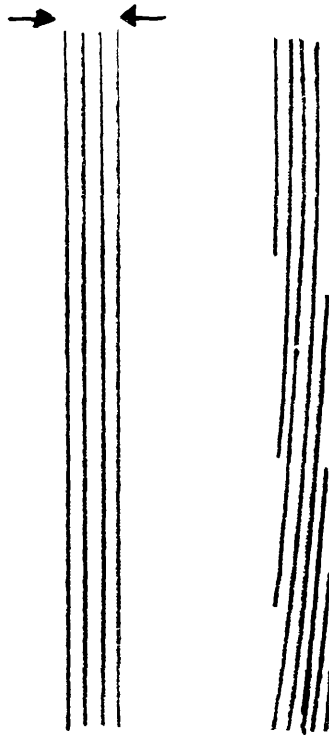


Figure 1.8

Qualitative representation of the interference pattern obtained in the biprism experiments of Chambers. The diagram on the left shows the fringe distribution without an applied flux whilst the diagram on the right shows the fringe distribution when a magnetised whisker was placed in the shadow region of the biprism.

However, it was Chambers [1960] who performed the first experiments demonstrating the quantum effects of an enclosed flux. The apparatus used was similar to the electrostatic biprism of Mollenstedt and Duker, but in order to introduce a magnetic flux between the beam trajectories an iron whisker was placed in the shadow region behind the central fibre of the biprism.

Such whiskers were known to be single magnetic domains and to taper with a slope of about 10^{-3} (figure 1.7). The taper gave rise to a radial leakage field which reduced the z-component of the flux. The z-component of the flux produced the enclosed flux which was responsible for the displacement of the interference fringes, but since this flux changed along the length of the whisker, so did the displacement of the fringes. In particular, the rate at which the flux was changing at any point along the z-axis was proportional to the radial field at that point. The fringes were thus tilted with respect to the z-axis, the slope of which was determined by the rate of loss of the enclosed flux to the radial field. Figure 1.8 shows a qualitative reproduction of the results from the experiments of Chambers. All the observed effects were in good agreement with the theory of Aharonov and Bohm.

1.3.3 The Aharonov-Bohm effect in metals

The experiment outlined above was performed in a vacuum. In metals and semiconductors however, the situation is very different in that electrons scatter elastically (without loss of energy) due to lattice defects, impurity atoms, grain boundaries and can also scatter inelastically from phonons. In samples of large size (>several tens of microns) electron transport may be thought of as diffusive, that is, electrons follow random paths within the material constantly colliding with lattice impurities or phonons, and then accelerating between collisions. In this regime, transport may be described in terms of an average collision rate [Blakemore 1985]. Carriers responsible for current flow are generally those close to the Fermi surface, so that the mean free path l , between collisions can be written as $l = V_f \cdot t$ where V_f is the Fermi velocity and t , the mean free time between collisions. Typical mean free path lengths in metals are in the order of a few nanometres.

In 1981 Al'tshuler, Aronov and Spivak (hereafter referred to as AAS) considered the effects of a magnetic vector potential on the conductivity of a disordered normal-metal cylinder [Al'tshuler, Aronov and Spivak 1981]. The calculation, based on the diagrammatic evaluation of the Green's function for electrons close to the Fermi-level

showed that in the quasi-1D case (where the circumference and thickness of the cylinder is small compared to the elastic scattering length) the correction to the classical Boltzmann conductivity varied with the applied flux as $(1/2)h/e$; half the normal Aharonov-Bohm flux. It was also found that the modulation amplitude of the fluctuations decreased exponentially with increasing circumference: a result which is in agreement with a simpler theories of the interference effect.

Seeking to verify the AAS effect, Sharvin and Sharvin [1981] coated a quartz filament of about 1.5 micron diameter and 1cm length with a magnesium film (figure 1.9). The choice of metal was motivated by the need to cool the sample to extremely low temperatures (less than 1K) so that the mean free path was as long as possible whilst ensuring that the sample did not become superconducting. The results, which are shown in figure 1.10, clearly showed the predicted $h/2e$ resistance oscillation. This was the first of many experiments which showed that the crucial transport parameter was the *inelastic* scattering length and not the total of inelastic and elastic scattering lengths. The present view is that although elastic scattering changes the phase of the wavefunction, it does so in a deterministic manner so that for a particular sample, a degree of correlation between trajectories is still possible. Any inelastic scattering process however, completely randomises the phase of the wavefunction so that no time-independent interference effects persist. Therefore, in using the word "coherent" it is usually understood that elastic collisions are to be allowed.

The identification that inelastic processes controlled quantum interference effects was an important step, as at low temperatures the inelastic scattering length can be several microns whereas the elastic scattering length is seldom more than a few Angstroms. Although the origin of the $h/2e$ oscillations may be formally understood from the diagrammatic calculations of AAS, the essential physics involved can be highlighted by considering a simple model of the interference process in which identical pairs of electron trajectories traverse the loop in opposite directions and interfere with each other back at their point of origin. The effect of this is to cause an enhancement of the wavefunction at the origin and increase the dwell time in this region resulting in an increase in the resistance of the sample. Gijs *et al* [1984] performed a detailed study of the AAS type resistance oscillations in cylindrical magnesium films in which the behaviour predicted by AAS was observed. This work also included a study of the temperature dependence of the oscillations, which supported the theory that the phase coherence length $l_\phi(T)$ decreases with increasing temperature.

Figure 1.9
Cylindrical Mg film used by
Sharvin and Sharvin to detect
AAS-type magneto-resistance
oscillations.
(Length=1cm , diameter=1.5 micron.)

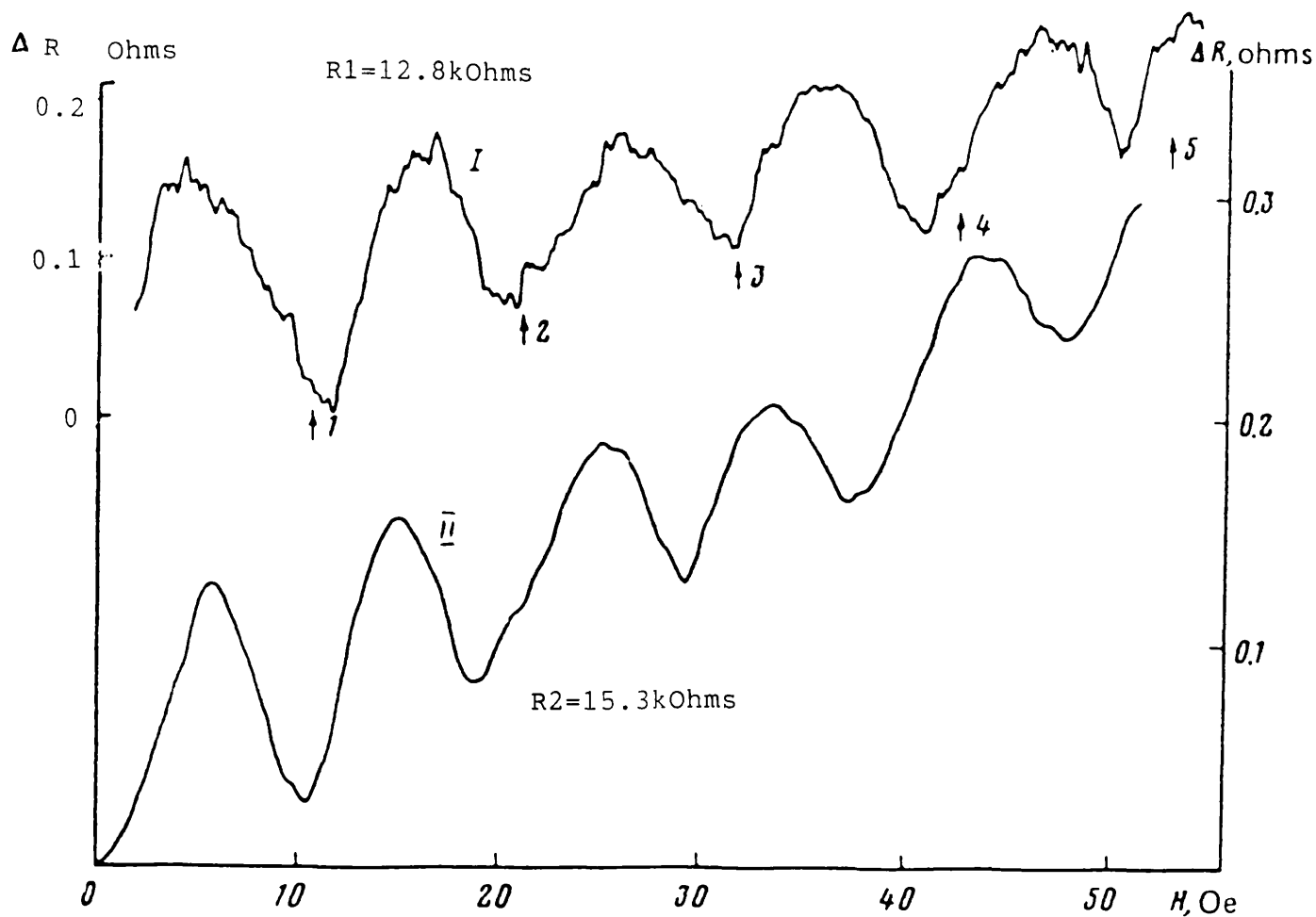
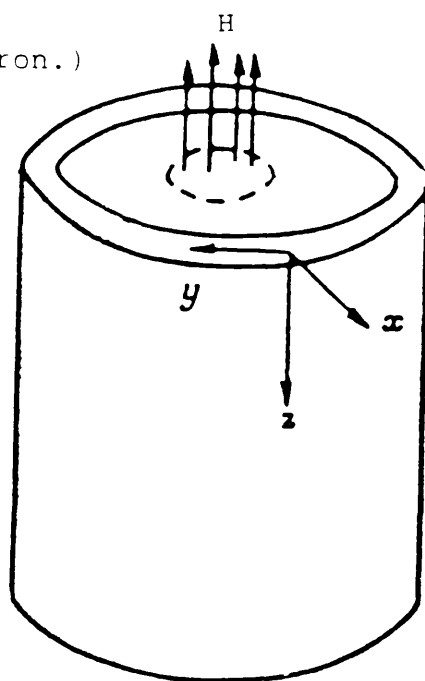


Figure 1.10

Magneto-resistance characteristic obtained by Sharvin and Sharvin for two different cylindrical Mg films. The arrows indicate the field corresponding to a whole number of $h/2e$ flux quanta in the cylinder cross-section.

Although flux-dependent oscillations had now been observed, there had been no clear observation of the AB conductance oscillation of period h/e and so research was initiated to search for these oscillations. Figure 1.11 shows a single loop of gold used by Umbach *et al* [1984] in their early experiments to detect conductance oscillations of period h/e . Figure 1.12 shows the magneto-resistance characteristic of this ring from which it is seen that there is no obvious structure in the plot such as h/e or $h/2e$ oscillations. On the contrary, the magneto-resistance spectrum appeared to consist of random noise, and it was only after careful examination that it was found that the data were, in fact, reproducible on cycling the magnetic field. Despite temperatures as low as 3mK and a ring diameter of 280nm with 45nm wide wires, the resistance fluctuations were only about 0.1% of the background resistance. The resistance fluctuations persisted up to fields of about 8 Tesla, which ruled-out oscillations arising from the AAS effect which were known to decay at much smaller fields (corresponding to fluxes of a few h/e). It was also found that the scale of the fluctuations increased with decreasing temperature approximately as $T^{-1/2}$. In these experiments however, it was not certain as to whether phase coherence had been maintained around the loop. In any case, an argument was put forward by Stone [1985] suggesting that even if phase coherence had been maintained, the variety of trajectories within the ring due to its low aspect ratio (the ratio between the ring diameter and the wire width) would result in apparently aperiodic fluctuations. The electrons traversing the inner and outer perimeters, for example, would acquire different phases from the field since the trajectories would enclose different amounts of flux. When the contributions from these trajectories were added together they would give rise to not just one periodicity in field but many, resulting in a confusion of signals at the output. The temperature dependence was explained in terms of an energy averaging of the incident electron distribution caused by thermal effects over an energy range $E = k_B T$. Assuming the existence of some energy correlation range E_C , over which the structure in the magneto-resistance data remains the same, the ring was thought of as averaging $n = k_B T / E_C$ uncorrelated channels. This would mean that if the signal in each channel was treated as a random variable, the relative size of the fluctuations would decrease in amplitude by a factor $n^{-1/2}$. Thus under the conditions where $k_B T \gg E_C$ (*i.e* many channels), this would explain the $T^{-1/2}$ temperature dependence.

Stone [1985] supported his findings by using a numerical simulation, based on a nearest-neighbour tight-binding model developed earlier by Lee and Fisher [1981]. The aperiodic fluctuations in the data hindered the search for the expected h/e oscillations and it was suggested that a necessary condition for observing unambiguous h/e oscillations was that the wire width be much smaller than the ring diameter.

With the intent of observing clear h/e oscillations of a single period, Webb *et al* [1985] manufactured a larger device (diameter 784nm) whilst keeping the wires at approximately the same width (41nm). This strategy was well rewarded, resulting in the first clear observation of h/e oscillations in single loops (figure 1.13). The oscillation period was in good agreement with that expected from the average area enclosed by the loop which was measured with an electron microscope.

1.3.4 The Aharonov-Bohm effect in semiconducting structures.

Other workers then turned to semiconducting structures. The motivation for using semiconducting structures, such as AlGaAs were as follows:-

1. The DeBroglie wavelength is typically much longer than in metals, so that the number of transverse modes in any structure would be substantially reduced, leading to enhanced interference effects.
2. Mobilities in such materials, most notably at a $\text{GaAs}/\text{AlGaAs}$ heterojunction, were known to be very high and offered near-ballistic transport in realistically-small devices.
3. The methods for producing ultra-thin layers with good integrity and the associated lithographic processes were both well advanced.

Bandyopadhyay *et al* [1986] performed some preliminary experiments based on the electrostatic version of the AB effect using a semiconducting structure shown in figure 1.14. Oscillations near the h/e period were found, although they were still only about 0.08% of the background resistance. Furthermore, only a few oscillation periods were seen, which although the period seemed to be correct, raised doubts as to whether they were in fact AAS $h/2e$ oscillations that were being damped by the field. Timp *et al* [1987] searched for the AB effect in a patterned-gate high-mobility $\text{GaAs}/\text{AlGaAs}$ heterostructure.

Figure 1.11

Gold loop fabricated by Umbach et al to search for magneto-resistance oscillations of period h/e . The loop has an inside diameter of 280nm and the wires are about 45nm in width.

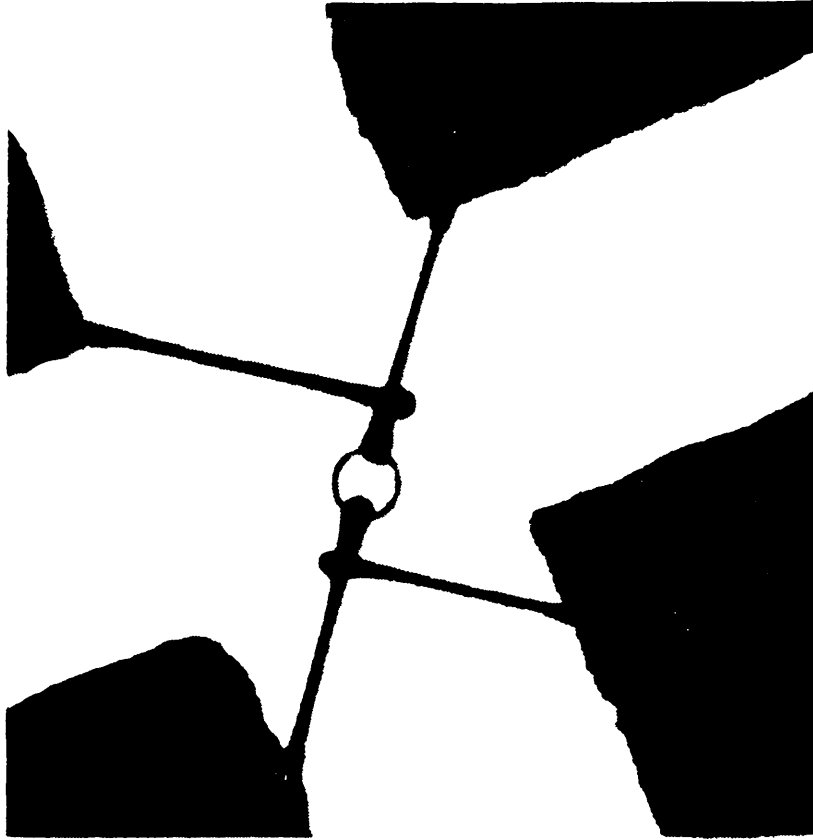
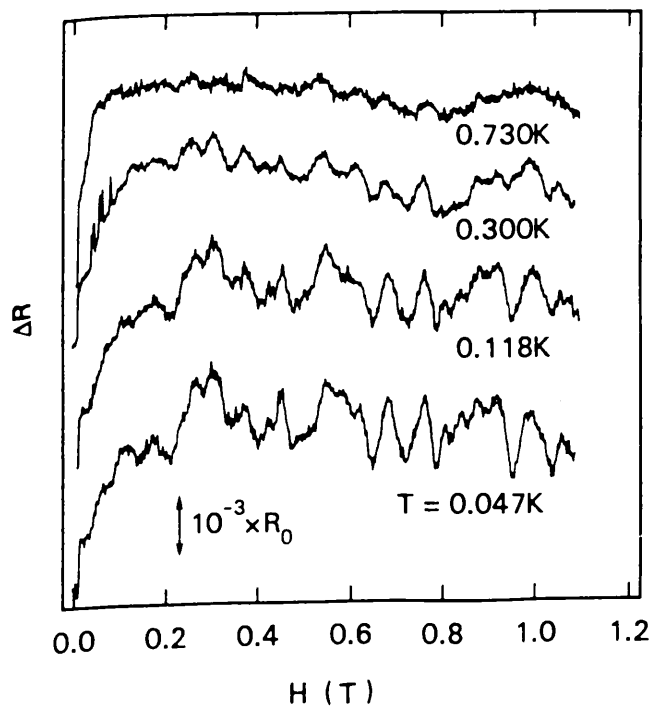


Figure 1.12

Magneto-resistance characteristics of the ring fabricated by Umbach et al. The zero-field resistance of the ring, R_0 , was 7.7 Ohms.



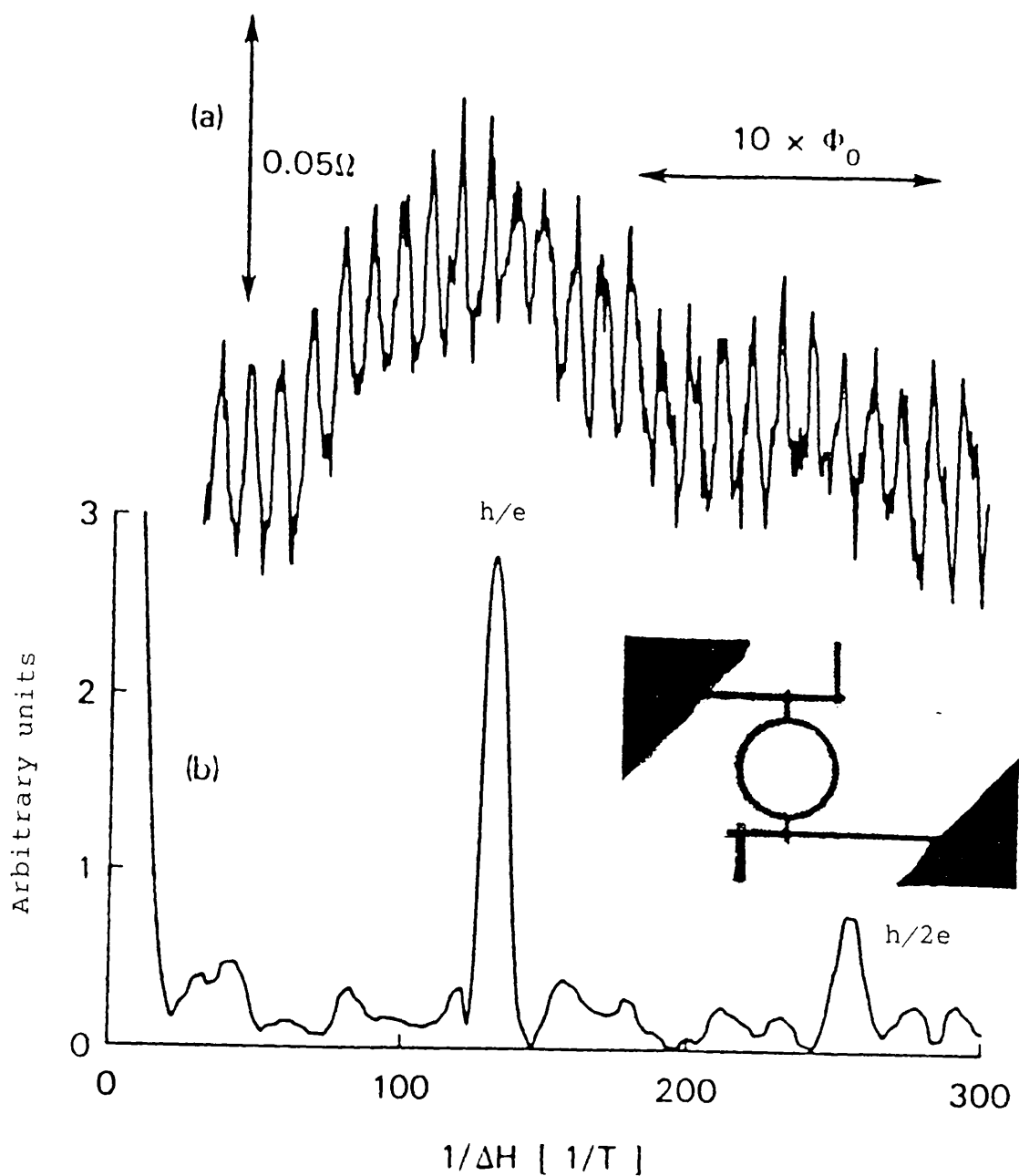


Figure 1.13

Clear observation of h/e magneto-resistance oscillations by Webb et al.

(a) Magneto-resistance characteristic

(b) Fourier power spectrum showing peaks corresponding to fluxes of h/e and $h/2e$. The large increase near the origin is due to aperiodic fluctuations.

The inset shows the ring used for the measurements. The inside diameter was 784nm and the wire width was about 41nm.

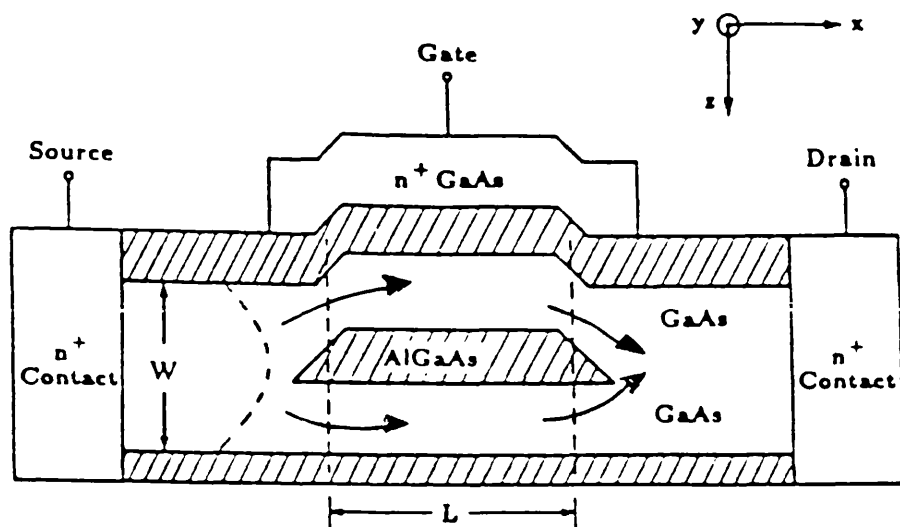


Figure 1.14

Structure of a device utilising the electrostatic Aharonov-Bohm effect proposed by Bandyopadhyay et al.

This type of structure had also been used by other workers [Ford *et al* 1987] and will be discussed in greater depth in chapter 4. The oscillation in the ring of Timp *et al* were much larger, being about 10% of the background resistance and were unambiguously identified (given the ring dimensions) as being of the h/e type. One of the other findings was that at around 0.3 Tesla the oscillations started to decay. The onset of this decay apparently occurred when the Larmor radius was half the wire width; an effect which will be discussed in greater detail in chapters 8 and 9.

1.4 The characteristics of different types of quantum interference processes.

Early experiments on metal rings showed a variety of conductance fluctuations with different periods of oscillation and different behaviours in strong magnetic fields. The $h/2e$ AAS effect is due to time-reversed pairs of trajectories encircling the ring and returning to the origin and is insensitive to ensemble averaging because the initial phase of the pairs is unimportant. These oscillations, which are typically no larger than the h/e contribution, do not self-average to zero.

The h/e trajectories are those which travel independently round each half of the ring and recombine at the other side of the ring. These trajectories start with an arbitrary phase relative to each other and therefore the h/e contribution self averages to zero as the number of trajectories is increased. If a magnetic field is uniformly applied to the wires in addition to the centre of the ring, the AAS trajectories now enclose differing amounts of flux, leading to different periods of oscillation. When these contributions are summed over all paths, the total oscillation amplitude is diminished. To illustrate these points consider the h/e contribution to be the sum of cosine terms over all paths with a period $\Phi_0=h/e$ and an arbitrary relative phase γ_p for each path.

If B_0 is the magnetic field required to produce a flux of Φ_0 through the centre of the ring then the contribution to the conductance from this process will vary approximately as:-

$$\Delta G(B) = \sum \cos\{(B/B_0 + \gamma_p).2\pi\} \quad (\text{Eqn 1.2})$$

where γ_p follows a random distribution over the paths. The AAS effect near zero field has the form:-

$$\Delta G(B) = -\sum \cos\{(2B/B_0).2\pi\} \quad (B \ll B_0) \quad (\text{Eqn 1.3})$$

and does not depend on any relative phase between paths. However, due to the fact that different trajectories will enclose slightly different areas, the value of B_0 will vary from one path to another in both h/e and $h/2e$ cases.

For the h/e contribution the conductivity correction can be written as:-

$$\Delta G(B) = \sum_p \cos\{(B/(B_0(P)) + \gamma_p).2\pi\} \quad (\text{Eqn 1.4})$$

Since the areas enclosed by the paths vary randomly within some range, this can be rewritten as:-

$$\Delta G(B) = \sum_p \cos\{(B/B_0 + \gamma'_p).2\pi\} \quad (\text{Eqn 1.5})$$

Where the difference between $B/(B_0(P))$ and B/B_0 has been absorbed into the phase factor γ'_p which is different to γ_p but still follows a random distribution.

Applying the same procedure to the AAS contribution gives:-

$$\begin{aligned} \Delta G(B) &= -\sum_p \cos\{(2B/(B_0(P))).2\pi\} \\ &= -\sum_p \cos\{(2B/B_0 + \gamma_p).2\pi\} \end{aligned} \quad (\text{Eqn 1.6})$$

Thus for fields of the order B_0 or greater, the AAS effect will start to average to zero like the h/e paths. The only effect the flux has on the h/e paths, which are more numerous, is to effectively change the random relative phases from γ_p to γ'_p , which is no different than the situation for which $B=0$. Thus, if h/e type oscillations are seen at all for $B=0$, the application of a magnetic field will not substantially alter the h/e contribution.

In addition to the periodic h/e and $h/2e$ types of oscillation, there are other fluctuations in conductance present that occur in rings or wires which show that these latter fluctuations are not strongly related to the topology of the system in which they are occurring.

A sharp decrease in magneto-resistance can be observed just above zero field in rings or wires. This effect is known as weak localisation and is due to the coherent backscattering of pairs of trajectories back to the origin. Weak localisation can be thought of as a more disordered form of the AAS effect in rings, but unlike the AAS effect, the areas enclosed by pairs of trajectories are defined by elastic collisions with impurity sites and not by the conductors forming the ring. Weak localisation effects therefore do not exhibit a single oscillation period but a superposition of periods relating to a range of areas enclosed by the trajectories (up to a maximum cut-off defined by the inelastic scattering length). Like the AAS effect, the localisation effect is quickly destroyed by the application of a magnetic field of the order of a few times the h/e flux quantum.

Another type of interference effect causing conductance oscillations has been reported to exist at fields of up to 11.5 Tesla [Leadbeater *et al* 1987]. These oscillations are also due to scattering from impurity sites and are believed to be related to the h/e type process, which is consistent with the persistence of the fluctuations to high fields. An analysis of this process by Lee and Stone [1985] has shown the magnitude of the fluctuations to be of the order of h/e^2 and independent of

sample size provided the wavefunction remained coherent within the sample. These oscillations were termed universal conductance fluctuations since it was also found that the magnitude of the fluctuations were independent of the degree of disorder and were only weakly dependent on sample shape. Figure 1.15 summarises the different types of interference processes and the types of electron trajectories which occur in each process.

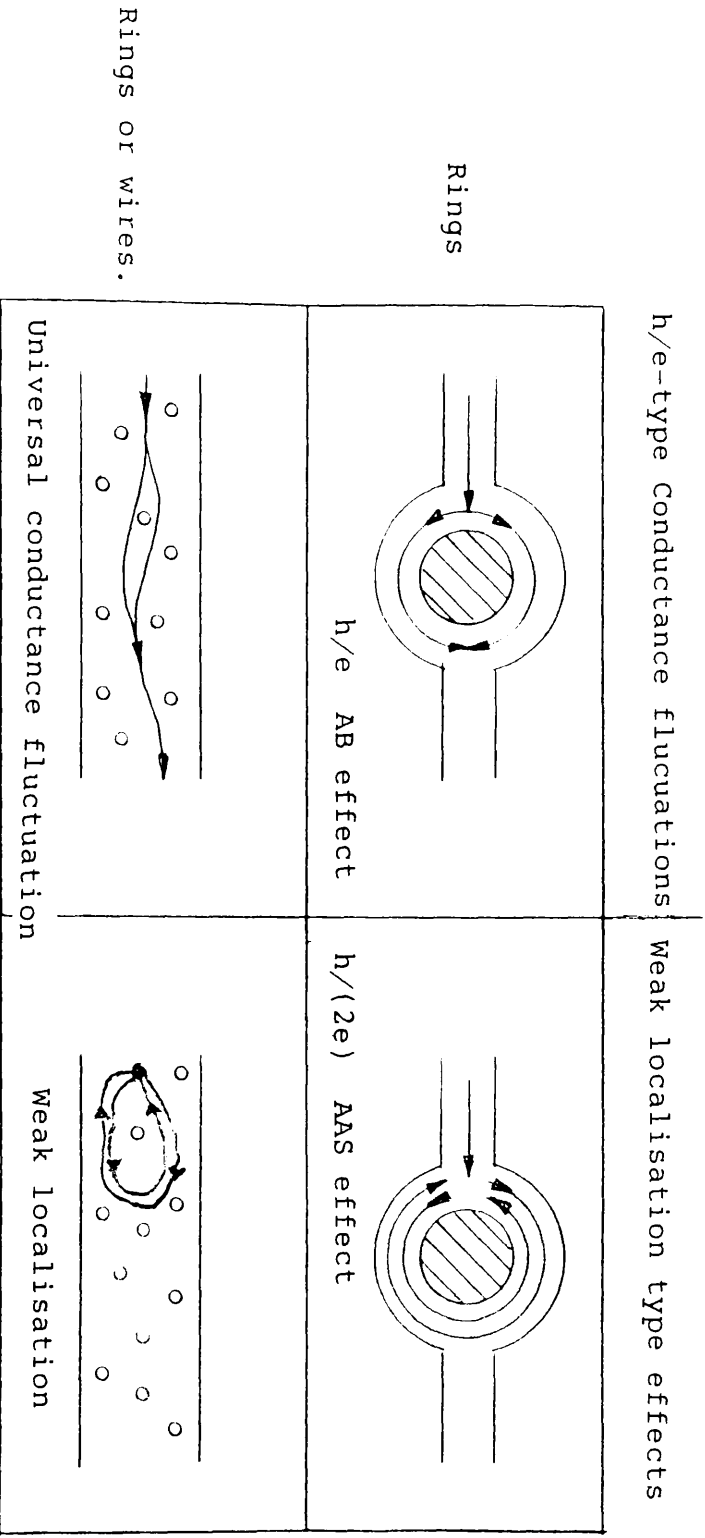
Universal conductance fluctuations in n-doped GaAs wires and split-gate heterostructures have been studied by Leadbeater and Taylor [Leadbeater *et al* 1987]. Figure 1.16 shows the measured magneto-resistance for a wire of 50nm thickness, 90nm width and 10 microns length. The four curves correspond to the magneto-resistances of the wire at temperatures of 4.2K, 13.4K, 26.0K and 51.5K from top to bottom.

The sharp decrease in resistance near zero field for each trace is due to the destruction of the weak localisation back-scattering effect with increasing magnetic field. Two main points are to be noted from these data. The first is that the amplitudes of the fluctuations decrease with increasing temperature, indicating an overall decrease in coherent backscattering due to the increase in phase-breaking collision

events. The second point is that there is a good correlation between the positions of the peaks in the traces at each temperature which were reported to be reproducible for temperatures below about 100K. This suggests that the microscopic defect configuration of the sample remains fixed at temperatures below 100K as the displacement of only one scattering site would be statistically indistinguishable from an entirely different random configuration, and can be expected therefore to give very different structure in the magneto-resistance characteristics [Cahay *et al* 1988]. The change in magneto-resistance structure above 100K is probably due to the thermal activation of electrons into defect traps such as the "DX-centre" [Iwata *et al* 1986] which would change the microscopic arrangement of scattering sites. Research is currently in progress to obtain more information concerning the precise mechanism [Taylor 1988].

Figure 1.17 shows cosine Fourier transforms of the full magneto-resistance curves (*ie* including negative field values). For single ring AB-type oscillations, the Fourier spectrum would normally show a distinctive peak around the oscillation frequency corresponding to a flux of h/e penetrating the ring. In the case of the universal conductance fluctuations, the spectrum of flux-enclosing areas can be revealed by

Figure 1.15 Electron trajectories leading to conductance fluctuations in rings & wires.



Figures 1.16 and 1.17 showing the magneto-resistance (MR) and Fourier transform of the MR respectively of a 90nm-wide wire. The four curves correspond to measurements taken at temperatures of 4.2K, 13.4K, 26.0K and 51.5K from top to bottom.

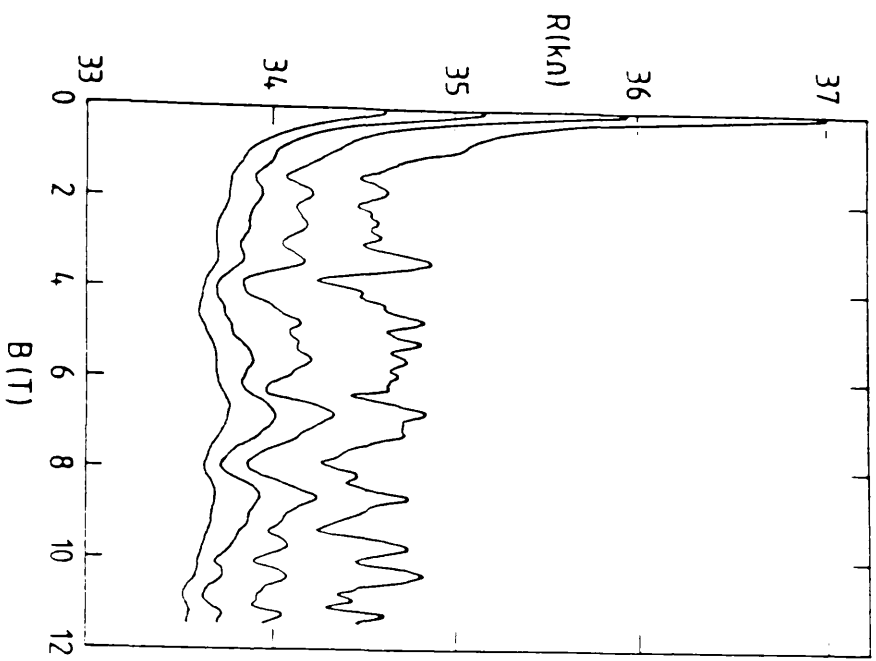


Figure 1.16

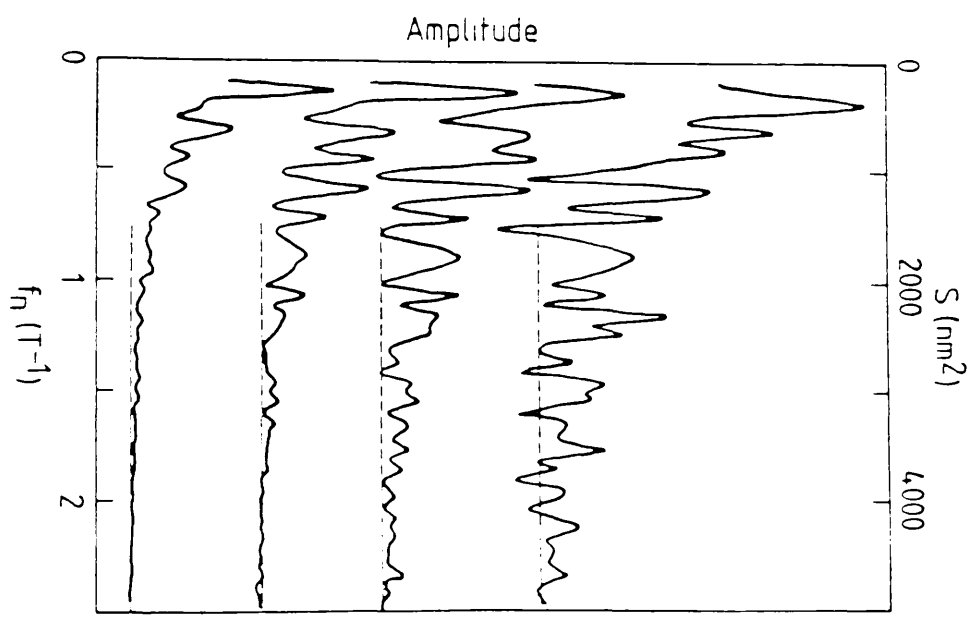


Figure 1.17

considering the contribution from each loop to be oscillating with a flux period of h/e . Thus, the peaks in the Fourier transform data can be identified with areas enclosed by contributing interference loops. The area scale is shown on the upper horizontal axis.

The oscillations in the Fourier transform data decay as the corresponding loop area increases due to the elimination of the larger loops from the interference process which is a result of the increased probability of a phase-breaking collision occurring during transit around the loop. This conclusion is substantiated by sharper decreases in the Fourier amplitude as the temperature was increased.

1.5 Summary

A review of quantum-mechanical effects in semiconductors was given and it was shown how the bulk transport properties could be altered as a direct result of these effects.

The different types of devices are classified by reference to their dimensionalities and the HEMT structure was shown to be a particularly versatile structure and will be discussed further in chapter 4.

The Aharonov-Bohm effect in vacuo, metals and semiconductors was then discussed. In vacuo, electrons move ballistically and produce clear interference fringes in electrostatic biprism experiments. In metals, the situation is quite different. The electrons can suffer many elastic collisions before interfering and so the transport can be thought of as diffusive. The transport in semiconductors is intermediate between these two extremes, although low temperature transport in the 2DEG in high mobility HEMT structures can be regarded as ballistic or near-ballistic over typical device lengths.

In metals and semiconductors a variety of conductance fluctuations are observed. These fluctuations are viewed as different aspects of the same fundamental process. Weak localisation is seen as a disordered form of the AAS effect and arising from the scattering of the electron trajectories from impurity sites whilst universal conductance fluctuations are seen as a disordered form of the h/e AB process caused by diffraction around impurity sites.

2. Quantum-mechanical modelling of devices in one dimension.

2.1 Introduction.

This chapter describes the quantum-mechanical modelling of 1D devices which was performed in the early stages of the project. Three numerical techniques for finding the quantum-mechanical transmission coefficient for a one-dimensional potential barrier are outlined with particular emphasis on an approach using transmission matrices.

The original motivation for concentrating on the transmission matrix method was that it was expected to provide a simpler way of calculating transmission coefficients for rectangular potential barriers and superlattices. It transpired that this was the case, but it was also found that the method provided a simple means of calculating the transit time through a system of barriers using the stationary phase approximation. In addition to these advantages, it was also recognised that, through the introduction of the 3-port scattering matrix, the theory could be extended to include multiply-connected ring structures, which are the subject of further study.

As a case study, the behaviour of the transmission coefficient and tunnelling time of a resonant tunnelling device was investigated using the transmission matrix technique.

2.2 Methods of calculating the transmission coefficient of a one-dimensional potential.

The transmission coefficient for an arbitrary one-dimensional potential cannot in general be found by simple analytic means. In cases where the potential is varying slowly and the incident electron energy is substantially less than the barrier height, the WKB approximation [Schiff 1955] can be used with some degree of success. However, in many important situations such as the double-barrier tunnelling problem, it has been shown [Collins 1986] that the WKB method is inadequate, especially in predicting the position of the resonant transmission. For these cases therefore, the WKB method cannot be relied upon and more robust numerical methods must be sought.

The first method considered was the method of Vigneron and Lambin [1980] who showed that if a region of arbitrary potential was considered to be between two regions of constant potential (at the end-points), the discretised Schrodinger equation could be efficiently solved by means of a continued fraction method. Once the problem has been discretised the method of Vigneron and Lambin solves the discretised Schrodinger equation exactly resulting in a terminating continued fraction. This method was used by Collins to find the transmission coefficients of single and double barrier structures, and has also been extended to find the complex transmission and reflection coefficients [Collins 1986] which contain information about the phase-shifts in the wavefunction.

Another method of solving for the transmission coefficient of a one-dimensional potential is by the use of the transmission matrix method and it is this method which is studied in this project. Its advantage over the continued fraction method is that considerably less computation is required if the form of the potential consists of rectangular-shaped barriers. It will also be shown that the complex transmission and reflection coefficients arise naturally from this method which enables a straightforward calculation of the tunnelling delay time through a system of barriers.

Finally, a more advanced approach using group-theoretical techniques is outlined. This approach is particularly well-suited to the calculation of the wavefunction through a system, although if only the transmission coefficient is required, the simpler transmission matrix method suffices.

2.3 Transmission matrix methods.

In the most usual formulation of transmission matrix theory one considers an arbitrary one-dimensional potential $V(x)$, such as that depicted in figure 2.1, which is constant outside the domain $-l < x < +l$. The barrier need not be centred on $x=0$, but it simplifies calculations if it is. The transmission matrix (T-matrix) then relates the coefficients of the forward and backward propagating plane waves on the right on the barrier to those on the left.

$$\mathbf{C} = \mathbf{T} \mathbf{A} \quad (\text{Eqns 2.1})$$

with

$$\mathbf{A} = \begin{pmatrix} A \\ A' \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} C \\ C' \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix}$$

Appendix A contains more information on the T-matrix and its properties, including its relation to an equivalent scattering matrix.

The utility of the transmission matrix can be demonstrated by considering the propagation of a wavetrain through a system of barriers similar to that shown in figure 2.2. If the individual transmissions $T_1, T_2, T_3 \dots T_n$ can be found then the total transmission is not the product of each of the individual transmissions. This is because in simply multiplying the transmissions, corrections due to multiple reflections within the structure are not taken into account. When the length of the system is comparable to the coherence length, multiple-reflection processes can be an important or even dominant factor in characterising the electrical behaviour of the structure. When the system is large compared to the coherence length, and there is no coherent interaction between the barriers, then the classical result outlined above is valid. In order to arrive at the correct formulation of the composite transmission matrix for a small system it is necessary to match the wavefunction at each interface. Since the transmission matrix relates the forward and backward propagating plane waves to the right of the barrier to those on the left by matching wavefunction and its derivatives at both interfaces, the correct rule for finding the composite T-matrix is to multiply the individual matrices for each of the barriers together.

However, in multiplying these matrices together, account must be taken of the null-potential region between the barriers, which will introduce an additional phase factor. The inclusion of these phase-shifts is essential in the modelling of quantum-mechanical systems as they are responsible for the energy resonances which occur in the system.

Care must be taken even if the barriers are in immediate contact with each other because the model calculation for the rectangular barrier assumed that the barrier was centred at $x=0$, which can only be true for one barrier in the system. To circumvent this difficulty, another matrix is used to connect the general solutions for each barrier and its effect in multiplying the current matrix product by it is to relocate the next barrier at the origin. Taking the centre as the reference point for each barrier, the general solutions are matched according to:-

$$A_{n+1} = B_r \exp(ik\Delta x) \quad (\text{Eqn 2.2})$$

$$A'_{n+1} = B'_n \exp(-ik\Delta x) \quad (\text{Eqn 2.3})$$

Where Δx is the distance between the centres of the barriers.

This transformation may be conveniently implemented using the iteration matrix, defined by:-

$$D = \begin{pmatrix} \exp(ik\Delta x) & 0 \\ 0 & \exp(-ik\Delta x) \end{pmatrix} \quad (\text{Eqn 2.4})$$

The general solutions are thus connected via:-

$$A_{n+1} = D_n \cdot C_n \quad (\text{Eqn 2.5})$$

Since A_n and C_n are related to each other by the T-matrix, a recurrence formula may be formulated:-

$$A_{n+1} = D_n \cdot (T_n \cdot A_n) \quad (\text{Eqn 2.6})$$

Figure 2.1

Type of potential considered in transmission-matrix calculations.

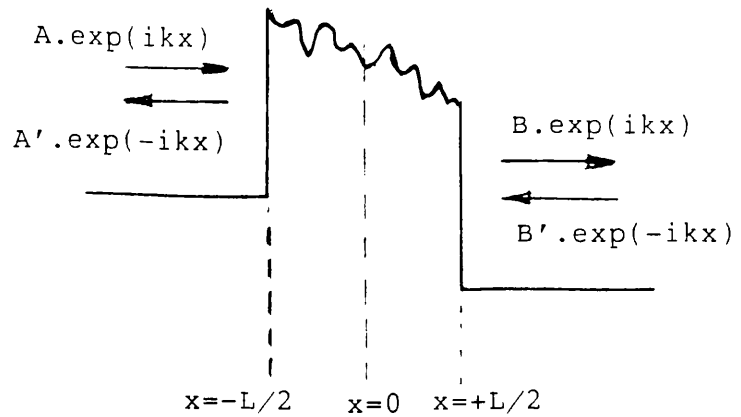


Figure 2.2

Type of potential which is well-suited for T-matrix techniques. The total transmission for the system is found by multiplying the T-matrices for each barrier and not their individual transmission coefficients.

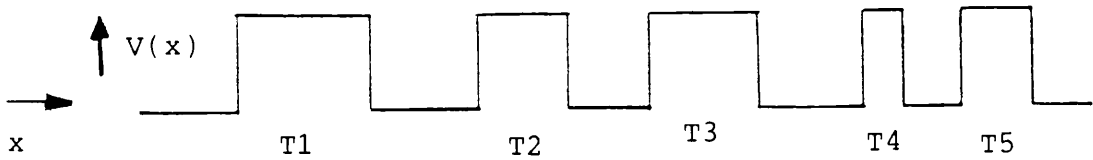
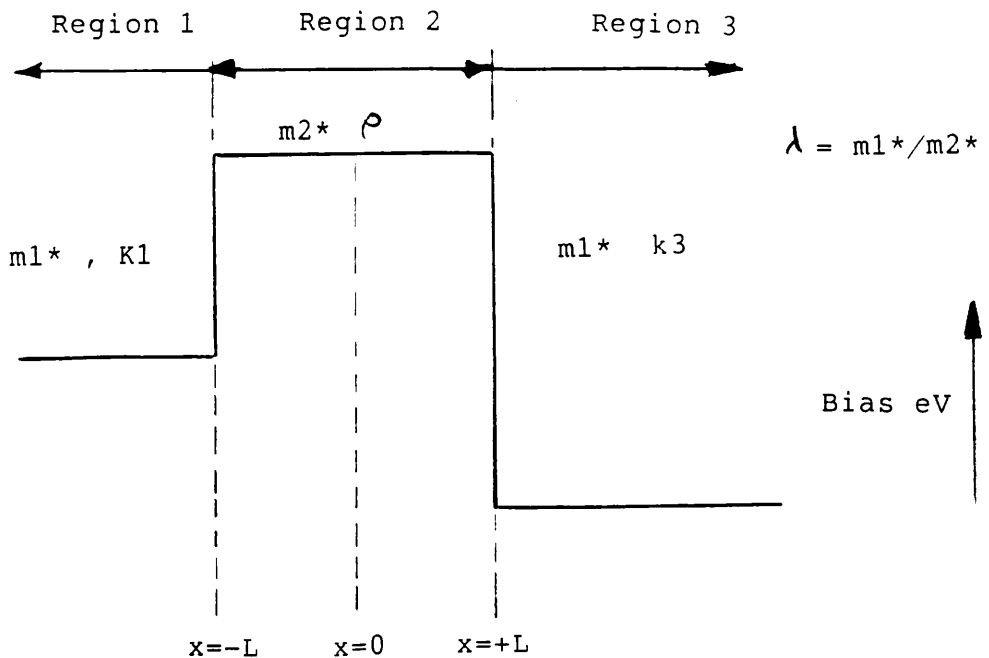


Figure 2.3 Model potential used to calculate T-matrix elements analytically.



and

$$\begin{aligned} C_{n+1} &= T_{n+1} (D_n (T_n A_n)) \\ &= (T_{n+1} D_n T_n) A_n \end{aligned} \quad (\text{Eqn 2.7})$$

This is now of the same form as equation (2.1), so that the quantity $(T_{n+1} D_n T_n)$ can be identified as the composite transmission for two barriers. For N barriers, this formula can be iterated to yield the total transmission matrix for the system.

$$T_{\text{tot}(N)} = T_N \Pi_{(N-1)}^1 D_n T_n \quad (\text{Eqn 2.8})$$

This iteration formula was then tested by calculating the transmission coefficient of a resonant double-barrier structure as a function of energy for which the results were well known.

2.4 The use of the transmission matrix to study a resonant tunnelling device.

The strategy for calculating the transmission of the double barrier device was to first calculate the T -matrix elements for a rectangular potential barrier depicted in figure 2.3 and then, by the use of the iteration matrix, multiply two identical matrices together to obtain the double-barrier result.

One convenient feature of the transmission matrix method is that differences in effective mass between the barrier and well regions could readily be taken into account. A calculation of the T -matrix elements (Appendix A) obtained by matching $\psi(x)$ and $1/m^*(x) d\psi(x)/dx$ at each interface yields the following results:-

$$\begin{aligned} T_{22} &= 1/2((k_1 + k_3)/k_1) \cosh(2\rho l) - i((k_1 k_3 - \lambda^2 \rho^2)/(\rho \lambda k_1)) \sinh(2\rho l) \\ &\quad \times \exp(i(k_3 + k_1)l) \end{aligned} \quad (\text{Eqn 2.9})$$

$$\begin{aligned} T_{21} &= -1/2((k_1 - k_3)/k_1) \cosh(2\rho l) - i((k_1 k_3 + \lambda^2 \rho^2)/(\rho \lambda k_1)) \sinh(2\rho l) \\ &\quad \times \exp(i(k_3 - k_1)l) \end{aligned} \quad (\text{Eqn.10})$$

where the energy of the incident wavetrain is considered to be less than the barrier height. The propagation constant, ρ , inside the barrier is thus:-

$$(2m^*/h^2)(V_{\text{barrier}} - E)^{1/2} \quad (\text{Eqn 2.11})$$

The parameter λ characterises the difference in effective mass in the barrier and well regions.

$$\lambda = m^*(\text{well}) / m^*(\text{barrier}) \quad (\text{Eqn 2.12})$$

Using these results and the iteration formula of equation 2.8, the transmission coefficient was calculated from the composite transmission matrix T_{tot} as $1/|T_{\text{tot}22}|^2$.

Figure 2.4 shows the transmission coefficient of a tunnelling device consisting of two barriers each 25Angstroms thick separated by 35Angstroms and 300meV high. The red curve is the transmission characteristic when the effective mass through the system was considered constant (*i.e.* $\lambda=1$) whilst the green curve shows the transmission when the effective mass in the well and barrier regions is different. To obtain an estimate of the mass appropriate for the barrier region, the following relations for the compositional dependence of the band-gap energy in $\text{Al}_x\text{Ga}_{(1-x)}\text{As}$ were used [Collins 1986].

$$E_{\Gamma_g}(x) = E_g(0) + 1.247x \quad x < 0.45 \quad (\text{Eqn 2.13})$$

$$E_{\Gamma_g}(x) = E_g(0) + 1.247x + 1.147(x-0.45)^2 \quad x \geq 0.45 \quad (\text{Eqn 2.14})$$

And the dependence of the effective mass on composition was taken as :-

$$m^*_{\Gamma}(x) = m^*_{\Gamma}(0) + 0.083m_{(\text{free})} \cdot x \quad (\text{Eqn 2.15})$$

To obtain the barrier height as a function of the composition parameter, x , a conduction/valence band offset must be assumed. Although some controversy still surrounds the value of the offset ratio, most authors seem to agree on a value of about 60:40 [Jaros 1986]. Taking this offset ratio, the effective mass in GaAs to be $0.067 \cdot m(\text{free})$ and the barrier height to be 300meV gives an effective mass in the barrier of 0.1 times the free electron mass.

It was found that increasing the barrier widths made the transmission peak sharper and the effect of increasing the width of the well was mainly to reduce the energy the resonance occurred at. The sharpening of the transmission peak is a result of a greater number of internal reflections taking place within the well resulting from the extra confinement caused by the mass discontinuity. The more partial waves exist inside the well region the greater the opportunity they have to add out of phase (thus destroying the resonance) unless they are very close to the quasi-bound energy.

The dependence of the resonance position with the energy can be understood by considering the double-barrier system to approximate the infinite potential well for which the energy levels are proportional to $1/(\text{width})^2$. An increase in the well width thus results in a lower resonant energy.

The results of the T-matrix calculations using a higher effective mass in the barrier region show that the resonant peak has been sharpened and shifted to a lower energy. The sharpening of the resonant peak can be attributed to the extra confinement of the wavefunction resulting from the increased effective mass. This has the same effect as increasing the barrier widths which can be seen by considering the exponentially decaying wavefunction inside the barrier which is of the form:-

$$\psi(x) = \exp(-\rho(m^*)x) \quad (\text{Eqn 2.16})$$

where ρ , the propagation constant, is dependent on the effective mass.

Therefore the larger the effective mass, the more rapidly the wavefunction is damped. Increasing the effective mass thus has the same effect as increasing the barrier thickness in the sense that the evanescent wave from the well is in both cases reduced by the time it reaches the outside of the barriers. It is the evanescent wave reaching the outside of the barriers which is responsible for allowing the trapped wavefunction to escape, so that a decrease in the evanescent wave amplitude means that the wavefunction inside the well has been more effectively confined.

The lowering of the resonant energy is a little more difficult to explain, as one might expect the resonant energy to increase as a result of the extra wavefunction confinement. To interpret this result, it is important to recognise that it is not the wavefunction which is now matched at the interfaces but the quantity $(1/m^*)d\psi/dx$. In the case where $m_2 > m_1$, this has the effect of making the slope of the wavefunction just inside the well lower than it would have been in the constant mass case. For a given wavefunction amplitude at the interface, the relation between the derivatives just inside the well in the two cases is:-

$$d\psi/dx(\text{discon}) = (m^*(\text{well})/m^*(\text{barrier}))^{1/2} d\psi/dx(\text{cont}) \quad (\text{Eqn 2.17})$$

By matching the evanescent waves inside the barrier to a sinusoidal form inside the well, it is seen that the lowering of the wavefunction derivative at the interface has the effect of causing the wavefunction to spread-out more inside the well. The wavefunction therefore behaves as if it were in a wider well, which is known to lead to lower bound energies. The introduction of a discontinuous mass between the barrier and well regions thus has the same qualitative effect as increasing the barrier and well width in the constant mass system.

The results for the constant mass case were checked for accuracy against the continued fraction method of Vigneron and Lambin using a very small mesh cell size of 0.25Angs. The two results agreed with each other to better than 0.025% for all energies up to 600meV apart from the resonant energy, where the agreement was about 0.4%. The computation time for 5000 energy points was 36s and 534s for the T-matrix and continued fraction method respectively using an HP9000 desktop computer.

TRANSMISSION AND TUNNEL DELAY TIME FOR A 25-35-25 DOUBLE BARRIER

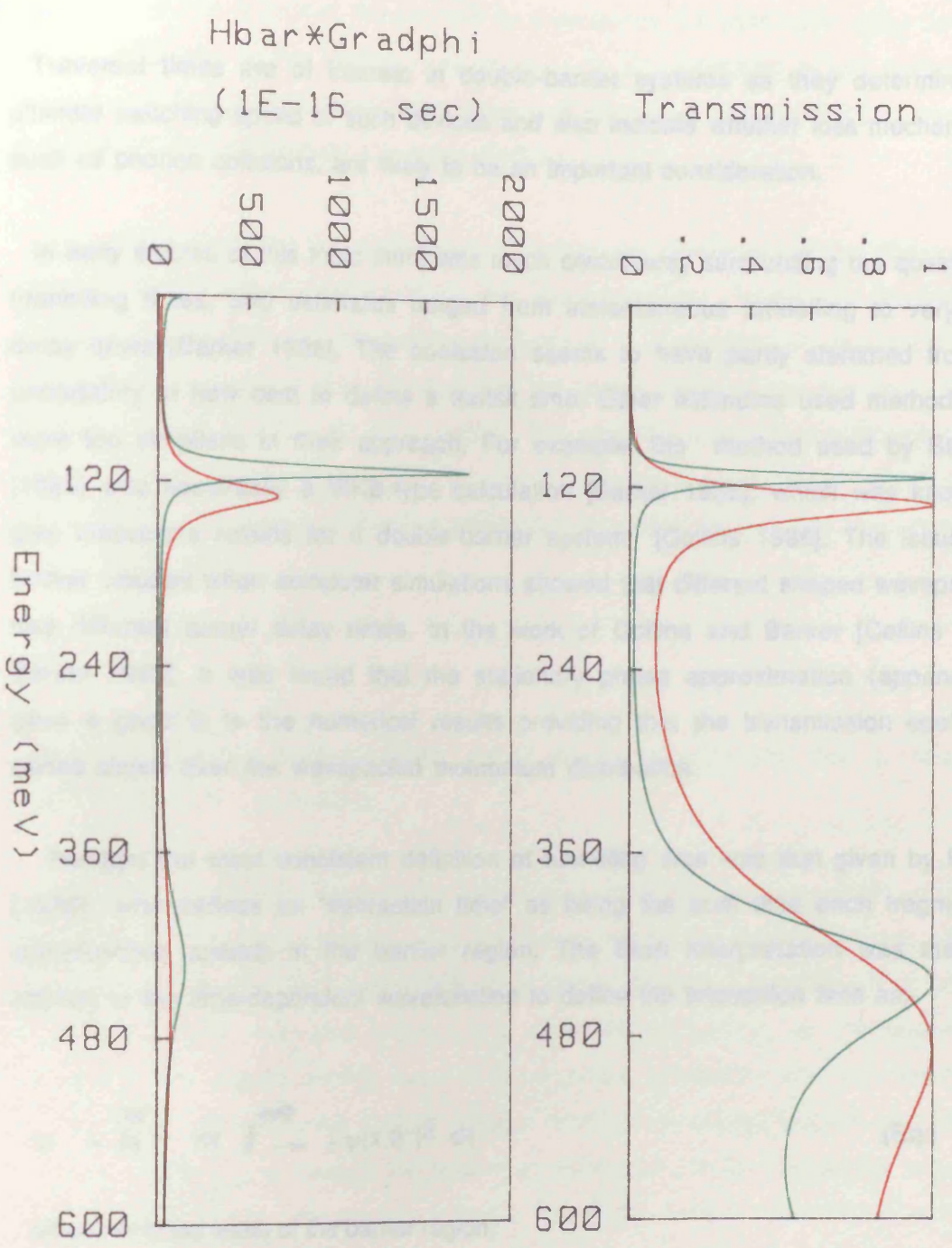


FIG 2.4

FIG 2.5

CONSTANT
MASS
VARIABLE
MASS

2.4.1 Traversal times.

Traversal times are of interest in double-barrier systems as they determine the ultimate switching speed of such devices and also indicate whether loss mechanisms, such as phonon collisions, are likely to be an important consideration.

In early studies of this topic there was much controversy surrounding the question of tunnelling times, and estimates ranged from instantaneous tunnelling to very long delay times [Barker 1986]. The confusion seems to have partly stemmed from an uncertainty of how best to define a transit time. Other estimates used methods that were too simplistic in their approach. For example, the method used by Stevens [1983] was essentially a WKB-type calculation [Barker 1985], which was known to give inaccurate results for a double-barrier system [Collins 1986]. The issue was further clouded when computer simulations showed that different shaped wavepackets had different tunnel delay times. In the work of Collins and Barker [Collins 1986, Barker 1985], it was found that the stationary phase approximation (appendix B) gave a good fit to the numerical results providing that the transmission coefficient varied slowly over the wavepacket momentum distribution.

Perhaps the most consistent definition of tunneling time was that given by Barker [1986] who defines an "interaction time" as being the sum time each fragment of wavefunction spends in the barrier region. The Born interpretation was therefore applied to the time-dependent wavefunction to define the interaction time as:-

$$\tau_i = \int_0^W dx \int_{-\infty}^{+\infty} |\psi(x,t)|^2 dt \quad (\text{Eqn 2.18})$$

where W is the width of the barrier region.

However, since the T-matrix method is a time-independent method and finds only the transmission coefficient and not $\psi(x,t)$, the stationary phase approximation will be used here to investigate the tunnel delay times for a double barrier structure as a function of energy.

Although the transit time for a continuous plane wave is a physically meaningless concept, one can approach the problem by considering the calculated delay times as being appropriate to the limiting case of an extremely broad wavepacket. However, as already mentioned these results will also serve as a good approximation in the case of more realistically sized wavepackets.

The calculation of the tunnel delay time involves the calculation of the phase of the complex transmission coefficient for the barrier system. Using the transmission matrix approach, this can be readily calculated as it is possible to calculate the real and imaginary parts of the wavefunction separately. This is not the case for the original method of Vigneron and Lambin. The complex reflection coefficient can be extracted without too much difficulty, but the complex transmission coefficient is more difficult to compute, although a satisfactory numerical method has been developed by Collins [1986].

2.4.2 Results of tunnelling delay time calculations for a resonant double-barrier potential.

Figure 2.5 shows the results of the tunnel delay time calculations plotted underneath the corresponding transmission plots for the cases of equal and different effective masses in the barrier and well regions. It is clearly seen that the delay time peaks at the under-barrier resonant energy and, to a lesser extent, at the over-barrier resonance. The constant-mass result is in excellent agreement with that obtained by Collins [1986] for the same system using a modified continued fraction method. The increase in delay time through the system at resonance can be interpreted in terms of the coherent waves undergoing multiple reflections before escaping in a manner analogous to the operation of a Fabry-Perot interferometer in the case of optical waves. The greater the confinement caused by the barriers, the larger the number of internal reflections the partial waves undergo before escaping and so the greater the tunnelling delay time. This also has the effect of sharpening the resonant peak in the transmission characteristic. This phenomenon can be described using the formula of Breit and Wigner [Landau & Lifshitz 1958, Messiah 1962] which was originally developed to investigate radioactive decay.

By specifying the width of the transmission peak as Γ , the lifetime of a particle trapped between the barriers is approximately:-

$$\tau = \hbar/\Gamma \quad (\text{Eqn 2.19})$$

It can be shown [Messiah 1962] that if the Breit-Wigner form of the transmission coefficient is used near resonance,

$$t(E) = 1/(E - E_r + i\Gamma) \quad (\text{Eqn 2.20})$$

then the probability density in the well decays with as time as:-

$$|\psi(t)|^2 = \exp(-2\Gamma/\hbar)t \quad (\text{Eqn 2.21})$$

with a time constant of $\hbar/2\Gamma$.

Furthermore the phaseshift in the wavefunction near resonance was found to be:-

$$\phi(E) = \text{Arctan}(\Gamma/(E - E_r)) \quad (\text{Eqn 2.22})$$

Thus from this formula it can be seen that the phase shift varies most rapidly when the energy is near resonance, which from the phase delay time result $\tau = \hbar d\phi/dE$, implies a long delay time near resonance. This analysis is therefore in qualitative agreement with the results obtained from the T-matrix calculations.

2.5 A group-theoretical approach to the transmission matrix method.

Peres [1983] has given a group-theoretical treatment of the transmission matrix technique. The Schrodinger equation was written as a first order multi-component equation by considering two functions, f and g , formed from linear combinations of ψ and $d\psi/dx$. The particular decomposition used was:-

$$f(x) = 1/2(\psi(x) + 1/(ik) \cdot d\psi(x)/dx) \quad (\text{Eqn 2.23})$$

$$g(x) = 1/2(\psi(x) - 1/(ik) \cdot d\psi(x)/dx) \quad (\text{Eqn 2.24})$$

$$\text{so that } \psi(x) = f(x) + g(x) \quad (\text{Eqn 2.25})$$

This decomposition can be seen as a generalisation for an arbitrary potential $V(x)$ of the null-potential result:-

$$\psi(x) = F. \exp(ikx) + G. \exp(-ikx) \quad (\text{Eqn 2.26})$$

so that for $V(x)=0$, the functions f and g are equal to $F.e^{ikx}$ and $G.e^{-ikx}$ respectively. The f and g functions can therefore be interpreted as the forward and backward propagating plane waves. In the case where $V(x)$ was not equal to zero, the f and g functions were interpreted in terms of forward and backward propagating waves inside an infinitesimally narrow "cut" in the potential profile.

The decomposition of the Schrodinger equation gave rise to two coupled first-order equations in f and g , but by the introduction of the two-component wavefunction (similar to a spinor, although no spin is considered in this theory) the equations could be reformulated as:-

$$i d\psi/dt = \{ (1-V(x)/2E) \sigma_z - V(x)/2E. i\sigma_y \} \psi, \quad \psi = (f,g)^T \quad (\text{Eqn 2.27})$$

where σ_y and σ_z are the standard Pauli matrices and the variable 't' is a length parameter through the system. The notation was chosen to indicate that the length variable could be thought of as being analogous to time in the normal form of the Schrodinger equation and the operator on the right hand side of equation 2.27 as being analogous to the Hamiltonian. Equation 2.27 can be rewritten in matrix form as:-

$$\psi(t) = \mathbf{A}(t) \psi(0) \quad (\text{Eqn 2.28})$$

where the 2 by 2 matrix $\mathbf{A}(t)$ can be thought of as a "time evolution operator".

From figure 2.1 it can be seen that $\mathbf{A}(t)$ is also the transmission matrix for the system expressing the forward and backward wave amplitudes either side of the system. The matrix $\mathbf{A}(t)$ can therefore be written as:-

$$\mathbf{A}(t) = \begin{pmatrix} T_{11} & T_{12}^* \\ T_{12} & T_{11}^* \end{pmatrix} \quad (\text{Eqn 2.29})$$

$$\text{with } |T_{11}|^2 - |T_{12}|^2 = 1 \quad (\text{Eqn 2.30})$$

(See Appendix A)

The transmission matrix therefore belongs to the $SU(1,1)$ group which again emphasises the role of the t -variable as a time-like parameter.

The time-like t -variable was written in equation 2.27 as $t=-kx$ to bring the equation into the same form as the time-dependent Schrodinger equation so that when t is going in the positive t -direction, x is going in the negative x -direction. This offers a rationale for the fact that the elements of the T -matrix for a potential barrier are most easily found by normalising the transmitted plane wave to unity; a condition which would at first sight seem more natural for the incident wave (Appendix A).

The group-theoretical method is then developed by writing a transmission matrix which is very close to the unit matrix as:-

$$\mathbf{A}_\delta = \mathbf{I} - \mathbf{S}_x \zeta - \mathbf{S}_y \eta + \mathbf{S}_t \tau \quad (\text{Eqn 2.31})$$

where \mathbf{I} is the unit matrix and,

$$\mathbf{S}_x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{S}_y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbf{S}_t = \frac{1}{2} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

(Eqns 2.32)

The coefficients ξ , η and τ are real and very small. Although one could conceivably work with this form of the transmission matrix by considering the transmission through a very narrow "slice" of the potential and multiply several such matrices together, the advantage of this approach is that the group properties of \mathbf{S}_x , \mathbf{S}_y and \mathbf{S}_t can now be used to generate a finite transfer matrix.

One can define a "vector" θ , to have components ξ , η and τ and write:-

$$\mathbf{A}_\delta = \mathbf{I} + \mathbf{S}.\theta \quad (\text{Eqn 2.33})$$

for equation 2.31.

Applying the infinitesimal transformation \mathbf{A}_δ to the wavefunction yields an infinitesimal change of:-

$$d\psi = d(\mathbf{S}.\theta) . \psi \quad (\text{Eqn 2.34})$$

so that the finite transformation matrix of equation 2.28 can be expanded as:-

$$\exp(\mathbf{S}.\theta) = 1 + (\mathbf{S}.\theta) + (\mathbf{S}.\theta)^2/2! + (\mathbf{S}.\theta)^3/3! + \dots \quad (\text{Eqn 2.35})$$

The group properties of the S-matrices may now be used to reduce the exponential expansion to:-

$$\exp(\mathbf{S}.\theta) = \text{Cos}(\theta/2) + \text{Sin}(\theta/2)(\mathbf{S}.\theta/\theta), \quad \theta = \xi^2 + \eta^2 - \tau^2 \quad (\text{Eqns 2.36})$$

which was compared to the form of equation 2.29 to find the T-matrix elements in terms of ξ , η and τ .

That the Schrodinger equation can be treated in this way is not especially novel, as a similar decomposition was used for the Klein-Gordon equation [Fresbach and Villars 1958]. What is novel in this case is that the group properties leading to the reduction of the exponential expansion in equation 2.35 can be realised by matrices of higher dimensionality. In particular, a three-dimensional representation of the $SU(1,1)$ group was used in which the complex expression $\mathbf{S}.\theta$ was replaced by a real one.

The vector $\mathbf{R}(X,Y,T)$ for the three-dimensional representation was defined in terms of the f and g components as follows:-

$$\begin{aligned}\mathbf{R} &= (X, Y, T) \\ X &= f^*g + fg^* \\ Y &= i(f^*g - fg^*) \\ T &= ff^* + gg^*\end{aligned}\tag{Eqns 2.37}$$

Once again the infinitesimal transform of equation 2.34 produces a rotation of the "vector" \mathbf{R} .

$$\mathbf{R}' = \mathbf{R} + (\mathbf{S} \cdot \theta) \mathbf{R}\tag{Eqn 2.38}$$

where $\mathbf{S} \cdot \theta$ is now a linear combination of real 3 by 3 matrices.

Current conservation then resulted in a restraint on the X, Y and T components of:-

$$T^2 - X^2 - Y^2 = 1\tag{Eqn 2.39}$$

The vector $\mathbf{R}(X,Y,T)$ is thus restricted to lie on the upper sheet of a unit hyperboloid so that the variation of the wavefunction in space was represented by a trajectory on the hyperboloid.

By introducing polar co-ordinates χ and ϕ to represent a point on the hyperbola (and therefore automatically taking into account the constraint of equation 2.39), the evaluation of the finite transformation matrix \mathbf{R} was formulated in terms of two real first order equations. The solution for χ and ϕ were then used to find the f and g functions and hence the wavefunction.

This method is useful in the numerical evaluation of the wavefunction because it imposes a restraint arising from the physical properties of the wavefunction (*ie* current continuity) and thereby acts to stabilise any numerical implementation. Otherwise, if the wavefunction is calculated directly from the T-matrix by considering the potential to be divided into a series of thin barriers, the absence of this constraint leads to numerical instability due to round-off errors.

In the work contained in this thesis, it was the transmission coefficients only which were calculated, for which a straightforward multiplication of the T-matrices was acceptable since there was no need to evaluate the wavefunction through the system and hence was not subject to the numerical instability described above. The group-theoretical method has been successfully applied to the calculation of wavefunctions in superlattices and disordered superlattices by Pepin [1987].

2.6 Summary.

In this chapter it was shown how the transmission coefficient of a one-dimensional potential could be calculated using transmission matrices. The method is particularly useful when the potential consists of rectangular-shaped barriers where it offers a considerable saving in computation time.

The other advantages of this method are that effective mass variations between barrier and well regions can be taken into account. Also, because the real and imaginary parts of the transmission coefficient arise naturally, it is a simple matter to calculate the tunnelling delay time within the stationary-phase approximation.

As a case study the transmission coefficient and tunnelling delay time of a resonant tunnelling device were calculated as functions of incident electron energy. The results for the case of constant effective mass in the barrier and well regions were found to be in excellent agreement with those obtained by Collins. The effect of different effective masses in barrier and well regions was to lower the resonant energy, make the resonance peak sharper and increase the tunnelling delay time. These results were attributed to an increase in confinement of the wavefunction between the barriers and at the same time a spreading of the wavefunction within the well region.

A more advanced approach using group-theoretical methods was outlined. This method, studied in detail by Pepin, is capable of accurately determining the wavefunction through a one-dimensional system. This accuracy and stability derives from the constraint imposed on the system of equations which arose from a need to ensure current continuity.

3. Theoretical origin of the Aharonov-Bohm effect.

3.1 Introduction.

In this chapter the Aharonov-Bohm effect is described in terms of the principle of gauge invariance. It is shown that overall gauge invariance of the Schrodinger equation is maintained if the change in the Hamiltonian (which is not gauge invariant) is accompanied by a change in the phase of the wavefunction.

Changes in the phase of the wavefunction have no effect on the expectation value of operators, but nevertheless lead to observable effects if two independent beams are allowed to interfere.

The phase change is related to the four-dimensional electromagnetic flux via a path integral over the trajectory of the particle. It is remarked that a phase change can also be achieved by the application of a transverse electric field.

The derived relations are then used in a simple model to determine the resulting wavefunction at the point where the two beams meet in both the magnetic and electrostatic case. An alternative derivation by Datta *et al* [1987] which considers the dispersion relation for two monomode wavepackets is outlined.

Finally, a theory by Berry [1984] is described. This theory can be considered to be a generalisation of the AB effect in other dynamical systems and offers another way of viewing quantum interference between two beams.

3.2 Gauge invariance.

The AB effect is a consequence of the gauge invariance of electromagnetic fields. This means that the differential of an arbitrary function $f(r,t)$ of space and time may be added to the electromagnetic *potentials* without altering the corresponding field strengths. For instance since

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (\text{Eqn 3.1})$$

letting $\mathbf{A} \rightarrow \mathbf{A} + \nabla f(r,t)$ will not have any effect on the strength of the magnetic field.

Maxwell's equation connecting the electric field in a vacuum to the time rate of change of the magnetic field can be satisfied if the electric field is written as:-

$$\mathbf{E}(r,t) = -d\mathbf{A}/dt - \nabla \phi(r,t) \quad (\text{Eqn 3.2})$$

so that the curl of $\mathbf{E}(r,t)$ will satisfy Maxwell's equation

$$\nabla \times \mathbf{E}(r,t) + \partial \mathbf{B}(r,t) / \partial t = 0 \quad (\text{Eqn 3.3})$$

Choosing $\mathbf{A}' = \mathbf{A} + \nabla f(r,t)$ means that the gauge-transformed scalar potential is $\phi' = \phi - df(r,t)/dt$ and hence the electric field is left unchanged.

Initially there appears to be a problem in applying these principles to the Schrodinger equation. From the analysis above it is seen that gauge transformation of the electromagnetic potentials had no effect on the field strengths and it is therefore to be expected that the dynamics of an electron moving in these new potentials will not be changed. However, the Hamiltonian for the new potentials is different, and therefore not gauge invariant with respect to the new potentials. In particular, the momentum operator transforms as:-

$$(\mathbf{P} - e\mathbf{A}) \rightarrow (\mathbf{P} - e\mathbf{A} - e\nabla f) \quad (\text{Eqn 3.4})$$

The Hamiltonian is thus gauge dependent whereas the energy is gauge invariant (up to a trivial constant).

It is not always the case that the Hamiltonian operator represents the energy of the system, although by the choice of an appropriate gauge function this can always be arranged [Kobe 1988]. Despite the gauge-dependent nature of the Hamiltonian, gauge invariance of the Schrodinger equation as a whole can be recovered by allowing the phase of the wavefunction to change.

$$\psi' = \exp(i \Lambda(r,t)) \psi \quad (\text{Eqn 3.5})$$

where $\Lambda(r,t)$ is an arbitrary function to be determined.

The change of phase of the wavefunction is not important when evaluating the expectation values of observables associated with an operator (L_0). Substitution of the new wavefunction into the standard expression for the expectation value of an observable (assuming $|\psi|^2$ normalised) yields:[†]

$$\begin{aligned} \langle L' \rangle &= \int \exp(-i\Lambda) \psi^* L' \exp(+i\Lambda) \psi d^3r \\ &= \int \psi^* L_0 \psi d^3r \\ &= \langle L_0 \rangle \end{aligned} \quad (\text{Eqn 3.6})$$

The expectation values of the operators are therefore left unchanged, which is consistent with the fact that the fields are unchanged by the gauge transform. Substituting $\mathbf{A} \rightarrow \mathbf{A} + \nabla f(r,t)$ and $\psi' = \exp(i \Lambda(r,t)) \psi$ into the time-dependent Schrodinger equation gives:-

$$\left\{ \frac{1}{2m^*} \cdot \left(\frac{\hbar}{i} \nabla - e\mathbf{A} - e \nabla f \right)^2 - e\phi \right\} e^{i\Lambda} \psi = i\hbar \frac{d}{dt}(e^{i\Lambda} \psi) \quad (\text{Eqn 3.7})$$

and upon using the results, (with $\gamma = -e\mathbf{A} - e \nabla f$)

$$(\nabla + \gamma) e^{i\Lambda} \psi = e^{i\Lambda} (\nabla \Lambda + \nabla + \gamma) \psi \quad (\text{Eqn 3.8})$$

$$\begin{aligned} (\nabla + \gamma)^2 e^{i\Lambda} \psi &= e^{i\Lambda} (\nabla \Lambda (\nabla \Lambda + \nabla + \gamma) + \nabla (\nabla \Lambda + \nabla + \gamma) \\ &\quad + \gamma (\nabla \Lambda + \nabla + \gamma)) \psi \end{aligned}$$

$$= e^{i\Lambda} (\nabla \Lambda + \nabla + \gamma)^2 \psi \quad (\text{Eqn 3.9})$$

[†] where L' is the gauge invariant form of L_0 .

The gauge-transformed Schrodinger equation is obtained:-

$$\frac{1}{2m^*} e^{i\Lambda} (\hbar \nabla \Lambda - e \mathbf{A} - e \nabla f + \hbar/i \nabla)^2 \psi - e \phi e^{i\Lambda} \psi = i\hbar (i d\Lambda/dt e^{i\Lambda} \psi + e^{i\Lambda} d\psi/dt) \quad (\text{Eqn 3.10})$$

The original (non-transformed) equation is thus recovered if

$$\hbar \nabla \Lambda - e \nabla f = 0 \quad , \quad \Lambda = (e/\hbar) f \quad (\text{Eqns 3.11})$$

and a new potential defined such that:-

$$-e\phi' = -e\phi + \hbar d\Lambda/dt \quad , \quad \phi' = \phi - df/dt \quad (\text{Eqns 3.12})$$

Gauge invariance of the Schrodinger equation as a whole is therefore restored if the phase of the wavefunction transforms as:-

$$\psi(r,t) \rightarrow \exp(ie f(r,t)/\hbar) \psi(r,t) \quad (\text{Eqn 3.13})$$

Of particular interest is a region in which there are potentials but no fields. In this case, the vector and scalar potentials can be represented by the differential of an arbitrary function g .

$$\mathbf{A} = \nabla g \quad (\text{Eqn 3.14})$$

$$\phi = - dg/dt \quad (\text{Eqn 3.15})$$

Letting $\mathbf{A}' = \mathbf{A} + \nabla f(r,t)$ and $\phi' = \phi - df/dt$ as before, it can be seen that the potentials in the field-free region can be completely eliminated by executing the gauge transform:-

$$f_\mu(r,t) = -g_\mu(r,t) \quad (\text{Eqn 3.16})$$

This enables one to use the simple free-particle Schrodinger equation and multiply any solution thus obtained by the phase factor $\exp(ie/\hbar \int \mathbf{A} \cdot d\mathbf{r} - \phi dt)$ to find the solution for a region with the potentials. Equations 3.14 , 3.15 and 3.16 give the gauge function required to do this as:-

$$f(r,t) = - \int \mathbf{A} \cdot d\mathbf{r} - \phi dt \quad (\text{Eqn 3.17})$$

It must be noted, however, that the boundary conditions for the problem have now been significantly altered. In particular, the quantity $f(r,t)$ is not single-valued under continuation around a closed circuit so that the usual boundary conditions of $\psi(\theta)=\psi(\theta+2\pi)$ are not observed [Olariu and Popescu 1985].

For a trajectory along a path P the integral

$$- \int \mathbf{A} \cdot d\mathbf{r} - \phi dt \quad (\text{Eqn 3.18})$$

is known as the four-dimensional electromagnetic flux.

In the generic AB experiment, a closed trajectory surrounds a region of non-zero field (figure 3.1) and $f(r,t)$ increases by an amount Φ for each circuit around the region.

$$\Phi = \oint_C (-\phi dt + \mathbf{A} \cdot d\mathbf{r}) \quad (\text{Eqn 3.19})$$

For closed circuits the quantum-mechanical effects of the potentials are therefore governed by the phase factor $\exp(-ie\Phi/\hbar)$ which is known as the non-integrable phase factor.

Equation 3.19 may be transformed using a four-dimensional version of Stokes' theorem into an integral over a surface spanning the path in four-dimensional space [Olariu and Popescu 1985].

$$\begin{aligned} \Phi = \int_{\text{surface}} & E_x dt dx + E_y dt dy + E_z dt dz \\ & + B_x dy dz + B_y dx dz + B_z dx dy \end{aligned} \quad (\text{Eqn 3.20})$$

This flux is different from the usual three-dimensional electric flux appearing in Coulomb's law of electrostatics. The above relation shows that an AB phase-shift can also be achieved by the application of an electric field. Other potentials can also give rise to interference effects and interference between neutron beams for example, has been observed [Colella *et al* 1975] as a result of differences in gravitational potential experienced by the two beams.

3.3 A simple analysis of the Aharonov-Bohm effect.

Using the concepts developed in the last section, the density $|\psi|^2$ at the point where the split beam is recombined is calculated (figure(3.1)). Taking the origin as the point where the beam was first split, the output wavefunction can be written as:-

$$\psi_{\text{out}} = \psi_1 \exp \left(i e / \hbar \int \mathbf{A} \cdot d\mathbf{r} \right) + \psi_2 \exp \left(i e / \hbar \int (-\mathbf{A}) \cdot d\mathbf{r} \right) \quad (\text{Eqn 3.21})$$

Assuming that the path lengths are equal and that the beam splits equally ψ_1 and ψ_2 may be written as:-

$$\psi_1 = \psi_{\text{in}} (1/2)^{1/2}, \quad \psi_2 = \psi_{\text{in}} (1/2)^{1/2} \quad (\text{Eqns 3.22})$$

and the direction of integration reversed for path 2 . This enables the two path integrals to be expressed in terms of a single line integral around a closed loop.

$$\int \mathbf{A} \cdot d\mathbf{r} + \int (-\mathbf{A}) \cdot (-d\mathbf{r}) = \oint \mathbf{A} \cdot d\mathbf{r} \quad (\text{Eqn 3.23})$$

By the use of Stokes' theorem, this can be written in terms of flux linkage Φ through the loop.

$$\oint \mathbf{A} \cdot d\mathbf{r} = \oint \mathbf{B} \cdot d\mathbf{s} = \Phi \quad (\text{Eqn 3.24})$$

The modulus-squared of the output wavefunction is thus:-

$$|\psi_{out}|^2 = (1/2) |\psi_{in}|^2 (e^{+ie\delta/\hbar} + e^{-ie\delta/\hbar})^2, \quad \delta = 1/2 \Phi$$

$$= (1/2) |\psi_{in}|^2 (1 + \cos(e \Phi / \hbar)) \quad (\text{Eqn 3.25})$$

The output wavefunction therefore varies with the flux linkage and has an oscillation period of h/e . This flux value is similar to the $h/2e$ flux quantum which arises in superconductivity, the $2e$ arising from the charge appropriate for a Cooper pair. There is however, an important difference between the origins of the two fluxes. In the AB experiment, the linking flux can take any value whereas the flux linkage through a superconducting ring is forcibly constrained to be an integer multiple of $h/2e$. It must also be remembered that electric charge must be conserved, so that when $|\psi_{out}|^2$ decreases, the wavefunction elsewhere must necessarily be enhanced. A more graphical demonstration of this will be given in chapter 8.

An alternative approach taken by Datta and Bandyopadhyay [1987] was to calculate the phase shift of a wavepacket travelling as a single transverse mode in a semiconductor AB ring (figure 1.14). The strategy was to calculate the dispersion relationships for the wavepackets in each of the two arms using the following Hamiltonian:-

$$H = (\mathbf{P} + e\mathbf{A})^2/2m^* + E_C(z) \quad (\text{Eqn 3.26})$$

Where $E_C(z)$ represents the energy profile of the conduction band in the z -direction.

Using a vector potential of the form,

$$\mathbf{A} = (B_y z, 0, 0) \quad (\text{Eqn 3.27})$$

the Hamiltonian expands to:-

$$H = 1/(2m^*) (\mathbf{P}_x^2 + eB_z \mathbf{P}_x + e\mathbf{P}_x B_z + B^2 z^2 + \mathbf{P}_y^2 + \mathbf{P}_z^2) + E_C(z) \quad (\text{Eqn 3.28})$$

The z-component terms are:-

$$E_z = \langle \psi | \mathbf{P}_z^2/2m^* + E_c(z) | \psi \rangle \quad (\text{Eqn 3.29})$$

which represents the energy of the sub-band caused by quantisation in the z-direction. The y-component terms simply express the normal parabolic E-k relation, representing free motion in the y-direction.

$$E_y = \langle \psi | \mathbf{P}_y^2/2m^* | \psi \rangle \quad (\text{Eqn 3.30})$$

The x-component of the energy can be expressed in terms of a modified \mathbf{P}_x operator by completing the square on

$$E_x = \langle \psi | 1/2m^*(\mathbf{P}_x^2 + 2eBz \mathbf{P}_x + B^2 z^2) | \psi \rangle \quad (\text{Eqn 3.31})$$

to yield

$$E_x = \langle \psi | 1/2m^*((\mathbf{P}_x + eBz)^2 + e^2 B^2 \langle z^2 \rangle) | \psi \rangle \quad (\text{Eqn 3.32})$$

Where $\langle z^2 \rangle$ is defined as the mean-square spread of the wavepacket:-

$$\langle z^2 \rangle = \langle \psi | z^2 | \psi \rangle - (\langle \psi | z | \psi \rangle)^2 \quad (\text{Eqn 3.33})$$

The total dispersion was found to be:-

$$E(k, B) = E_0(z) + \hbar^2 k^2 / 2m^* + (\hbar k_x + eB \langle z \rangle)^2 / 2m^* + e^2 B^2 \langle z^2 \rangle / 2m^* \quad (\text{Eqn 3.34})$$

where $E_0(z)$ is the energy of the bottom of the 1D sub-band.

It was then assumed that the energy of each of the eigenmodes in the two channels was equal and that the spread of the wavepackets in the two channels was equal.

Figure 3.1 Generic Aharonov-Bohm interference experiment.

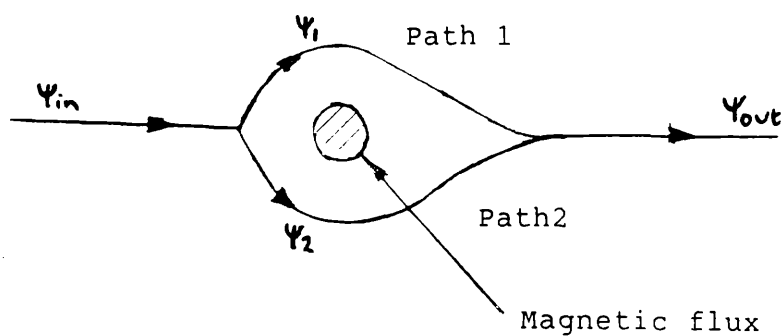
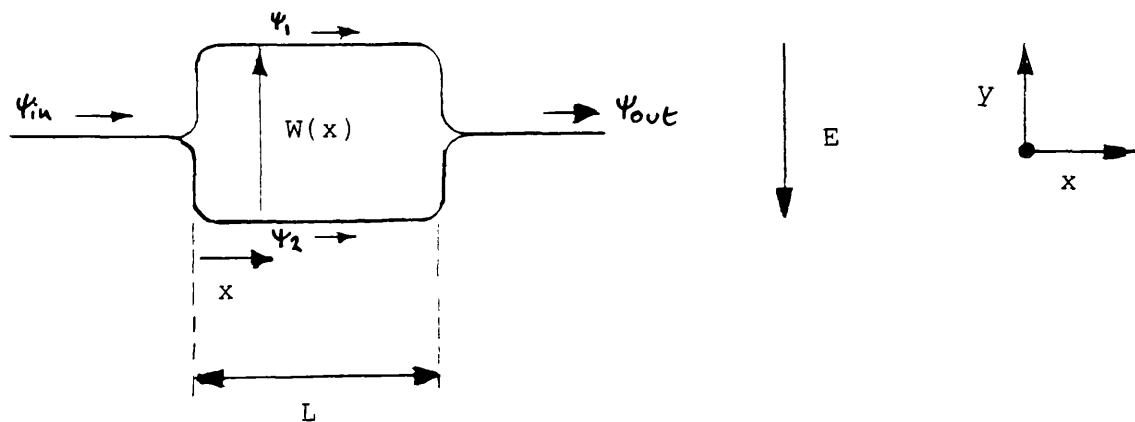


Figure 3.2 The electrostatic Aharonov-Bohm effect.



The upper and lower paths are held at different potentials by the application of a transverse electric field.

Taking the zero for the z co-ordinate as being in the centre of the paths and equating energies of the two wavepackets resulted in a solution for the difference between the two wavevectors K_1 and K_2 in the x-direction in the two channels.

$$(k_1 - k_2) = eB(z_1 - z_2)/\hbar + e(\phi_1 - \phi_2)/\hbar V_x \quad (\text{Eqn 3.35})$$

where $V_x = \hbar(k_1 + k_2)/2m^*$ and ϕ_1 and ϕ_2 are the electrostatic potentials of the two paths. Multiplying by the length of the channels, L , gave a value for the Aharonov-Bohm phase shift Φ .

$$\Phi = eBL(z_1 - z_2)/\hbar + e(\phi_1 - \phi_2)L/\hbar V_x \quad (\text{Eqn 3.36})$$

Which again shows that a transverse electric field in addition to a magnetic field can induce a phase difference between wavefunctions following two different paths.

3.4 The electric Aharonov-Bohm effect.

A simplified analysis along the same lines as the magnetic effect can be performed by considering an idealised ring with one-dimensional conductors, illustrated in figure 3.2.

The ring, of length L and width W , has its upper and lower arms held at potentials ϕ_1 and ϕ_2 respectively by the application of a transverse electric field. The difference in phase between the two plane waves travelling in the arms can be seen to arise from the electric part of the 4-dimensional electromagnetic flux described in section 3.2. Concentrating on the case for which $B=0$, $E_x=0$, and $E_z=0$ and taking (without loss of generality) the input wire to define the y co-ordinate datum point, then the relative phase shift between the beams is:-

$$\Delta S = \int (\phi(t)_{\text{upper}} - \phi(t)_{\text{lower}}) dt \quad (\text{Eqn 3.37})$$

Assuming the transverse electric field to be a constant, the potentials can be written as:-

$$\phi(t) = E_y (\pm W(x(t))/2) \quad (\text{Eqn 3.38})$$

so that

$$\Delta S = \int E_y W(x(t)) dt \quad (\text{Eqn 3.39})$$

An electric flux can thus be defined via the relation:-

$$\Phi_e = E_y \int W(x)/V_x(x) dx \quad (\text{Eqn 3.40})$$

although this is not in the usual form of a flux, since the associated area is not normal to the electric field and also involves the x-component of the velocity. If W is assumed constant and the electric field is small enough so that V_x can be assumed constant, the expression for the electric flux reduces to:-

$$\Phi_e = (E_y/V_x) W L \quad (\text{Eqn 3.41})$$

in which the area enclosed by the trajectories can be identified as WL , so that the actual electric flux is the quantity E_y/V_x .

As in the case of the magnetic effect, this phase-shift can be achieved without the field doing any net work on the wavepacket. This is because the electric field can, in principle, be confined to a region in which the arms of the ring run perpendicular to the field. In contrast to the magnetic effect however, the particle does experience fields as it passes through the regions of different potential. Nevertheless, energy is still conserved as it is in analogy to the zero bias tunnelling experiment.

Figure 3.3 shows a schematic diagram of a ring under the influence of a transverse electric field. Whilst the energy of the wavepacket is conserved overall, the partial wavepacket in the upper arm (in this case) can be reflected by the effective barrier caused by the electric field.

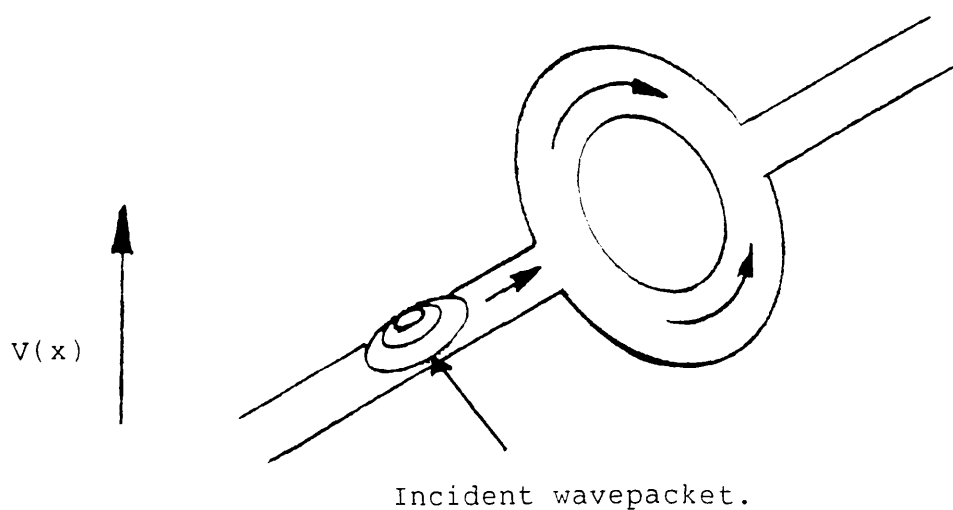


Figure 3.3

Wavepacket incident on an AB ring under the influence of a transverse electric field. The partial wavepacket traversing the upper arm encounters a potential barrier as a result of the field, whereas the wavepacket traversing the lower arm encounters a potential well.

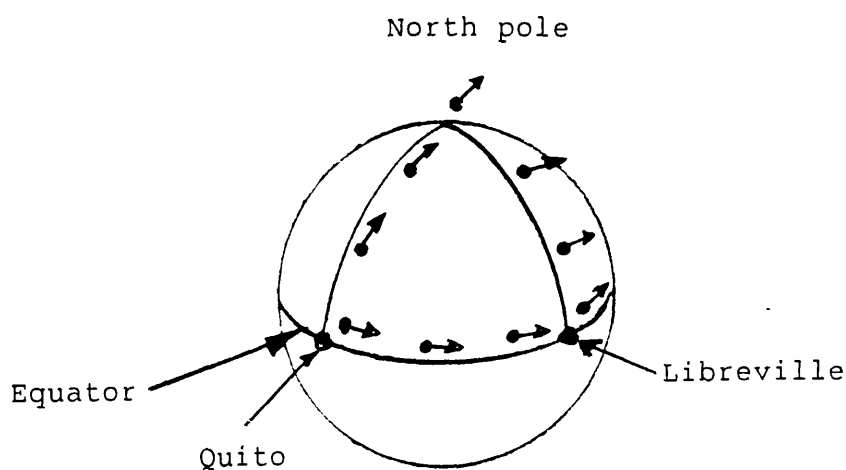


Figure 3.4

An example of Berry's geometrical phase arising from the parallel transport of a vector over a curved surface.

This means that although the phase-shifts can be arranged so that destructive interference would occur, the two partial wavepackets would have very different amplitudes at the output wire, resulting in a loss of modulation depth in the AB oscillations. This effect would become more severe as the device dimensions WL are reduced.

In structures similar to Datta's [1986], W is necessarily small to promote single-mode propagation of the wavepacket, leaving the length as the only adjustable parameter. The gate voltage in Datta's device of about 1mV would appear to be correct in comparison with the expected energy of about 10meV so that not too much reflection would have taken place.

3.5 Berry's geometric phase factor.

Berry has studied the transport of a quantal system around a closed circuit [Berry 1984]. In this analysis, the AB effect can be regarded as a special case of a much more general theory. It was stressed that in addition to the familiar dynamical phase factor $\exp(iEt/\hbar)$, there is also a circuit-dependent factor $\exp(i\gamma)$. Attention was focussed on the circuit-dependent contribution (which is not single-valued under continuation around the circuit) and its value was calculated in general terms as a function of the eigenstates of the Hamiltonian. The time evolution of the wavefunction for a system being transported around a path $\mathbf{R}(t)$ in parameter space was expressed as:-

$$| \Psi(t) \rangle = \exp \left\{ -i/\hbar \int_{t_0}^t dt' E_n(\mathbf{R}(t')) \right\} \cdot \exp(i \gamma_n(t)) \cdot | n(\mathbf{R}(t)) \rangle \quad (\text{Eqn 3.42})$$

The quantity $| n(\mathbf{R}(t)) \rangle$ is the instantaneous basis of eigenvectors satisfying the time-independent Schrodinger equation at any instant (the adiabatic approximation). The evolution of states $| n(\mathbf{R}(0)) \rangle$ with $\mathbf{H}(\mathbf{R})$ to states $| n(\mathbf{R}(t)) \rangle$ can therefore be viewed as the effects of the direct action of forces on the system trajectory.

$$\mathbf{H}(\mathbf{R}) | n(\mathbf{R}) \rangle = E_n(\mathbf{R}) | n(\mathbf{R}) \rangle \quad (\text{Eqn 3.43})$$

The first exponential term in equation 3.42 is the usual dynamic phase factor describing the time evolution of the eigenstates and the second exponential is the circuit-dependent phase factor.

For this equation to hold, the Hamiltonian must be changed slowly around the circuit so that the adiabatic approximation can hold. That is, the basis states of the wavefunction must be eigenvectors of the instantaneous Hamiltonian. By substitution of the trial wavefunction back into the time-dependent Schrodinger equation, Berry was able to show that the total phase change due to the circuit-dependent contribution was:-

$$\gamma_n(c) = i \int \langle n(\mathbf{R}) | \nabla_{\mathbf{R}} n(\mathbf{R}) \rangle \cdot d\mathbf{R} \quad (\text{Eqn 3.44})$$

Stokes' theorem then enabled a reduction to a surface integral of the form:-

$$\gamma_n(c) = - \int V_n(\mathbf{R}) \cdot d\mathbf{S} \quad (\text{Eqn 3.45})$$

Where V_n can be identified as a magnetic field in the case of the AB effect.

Berry's most significant contribution was that he showed that this "geometrical phase factor " could be observed in systems other than the Aharonov-Bohm experiment. Pursuing this line of thought Chiao and Wu [1986] asked whether one would be able to observe the geometric phase shift for a photon. A short time later, Tomita and Chiao [1986] carried out an experiment in which laser light of known polarisation was introduced into one end of an optical fibre wound into a helix which was then subjected to a magnetic field. The polarisation of the light emerging from the opposite end of the fibre was measured and found to be in very good agreement with that predicted from theory.

However, such geometrical phase factors can arise even in purely classical situations [Berry 1987]. Consider, for example, a journey around the world in which one carries a pointer (figure 3.4). Setting-off from Quito near the equator and travelling east to Libreville with the pointer pointing eastwards, one then turns north taking care not to change the direction of the pointer. At the North pole one then heads south to the starting point at Quito, again ensuring that the pointer is not turned. On arriving at Quito, one finds that the direction of the pointer is at right angles to its original direction: an effect that would not have been present if the Earth's surface had been flat.

This is an example of a geometric phase factor which is seen to arise from the curved surface over which the pointer moves. Berry's phase, and hence the AB effect, is thus a consequence of the curvature of parameter space in which the a particular system moves. The curvature in the case of the AB effect is caused by the magnetic vector potential. The case of a dynamic system undergoing such a journey in the example above, could be represented by a spinning weather vane. On completing the journey whilst spinning at a constant angular velocity ω , the *total* angle the vane would have turned through (with respect to its starting direction) would be :-

$$\Theta = \omega.T + \gamma_C \quad \text{(Eqn 3.46)}$$

Where $\omega.T$ is the dynamic phase, and γ_C the geometric phase ($\pi/2$ in this case).

3.6 Summary.

This chapter has discussed the underlying theory of the AB effect in which the phase difference between two interfering beams can be related to the four-dimensional electromagnetic flux enclosed by the beam trajectories. A phase difference between the beams can therefore be achieved both by the presence of a transverse electric field or by an enclosed magnetic flux.

Analysis of a simple model in the magnetic case showed that the probability density at the point where the two separate beams interfered was a simple sinusoidal function of the magnetic flux enclosed by the trajectories. The flux oscillation period was h/e and the probability density reached a minimum of zero at a flux value of $1/2(h/e)$.

The following chapters will consider the structures in which the AB effect can take place and also how the simple model can be improved upon.

4. Structures for achieving confinement of carriers in one and two dimensions.

4.1 Introduction.

In recent years there has been increasing interest in quantum-mechanical fluctuation phenomena in small semiconductor devices. As the length scales and dimensionality of these systems are reduced the fluctuation phenomena associated with the breakdown of ensemble averaging of random processes at a microscopic scale become more apparent.

Two experiments concerned with these phenomena are described. These experiments achieved the further confinement of carriers in the inversion layer of silicon MOSFET devices by the techniques of deep etching and electrostatic confinement respectively. Confinement of the 2DEG in HEMT structures is also considered and described in detail, as these structures have been configured by several workers to act as AB rings.

4.2 Transport phenomena in the inversion layers of silicon NMOS transistors.

4.2.1 General.

This section describes two approaches to the study of transport phenomena in the inversion layers of NMOS transistor structures.

The first approach considered is a multi-terminal MOSFET fabricated by Skocpol [1986] which enabled a detailed examination of the conduction processes in an FET. The second approach, used by Hartstein [1986] uses electrostatic confinement of the two-dimensional electron gas to form a quasi-1D channel in the inversion layer.

4.2.2 Transport in a multi-terminal MOSFET.

Figure 4.1 shows a schematic diagram of Skocpol's device. A NiCr mask was patterned onto a conventional MOSFET structure using electron beam lithography. The sample was then subjected to reactive ion etching to remove material not protected by the NiCr mask down to the lightly-doped silicon substrate. The object of this structure was that a current could be passed through the narrow channel formed in the inversion layer and the local potential probed at various points by connecting high-impedance voltmeters to the side-branches.

Figure 4.2 shows the resistance of a 0.1 micron wide by 0.25 micron long segment of channel between probes 2 and 3 as a function of gate voltage and temperature. The results clearly show jumps in the resistance of the segment due to single electron traps charging and discharging with temperature-dependent characteristic times. These traps are believed to exist only a few Angstroms from the SiO_2/Si interface and can have energies near the Fermi energy in the inversion layer. Once a trap has been charged, it acts as a scattering centre and can readily affect the resistance of devices of this size where they may be only in the order of a hundred other scatterers. The duty cycle of the switching is seen to vary as the energy of the trap is swept past the Fermi-energy due to the activation energy between the trap energy and the Fermi energy increasing. As the temperature is lowered, the thermally-activated switching rate is seen to decrease and only a single trap switching on but not off is observable at the lowest temperature of 5K.

The magneto-resistance measurements on the device revealed the presence of universal conductance oscillations whose rms amplitude was about $0.06e^2/h$ superimposed on an average background conductance of $1.1e^2/h$. It was known from the work of Lee and Stone [1985] that provided the wavefunction remained coherent within the sample, one could expect an r.m.s variation of approximately e^2/h if the microscopic arrangement of scatterers is changed or if approximately one quantum of flux was applied. Skocpol therefore interpreted the results in terms of many interconnected quantum subunits whose length was the estimated inelastic diffusion length. Each coherent subunit was considered to be contributing a random component of e^2/h to the overall conductance.

A quasi-1D channel therefore, can be considered to be composed of N ($N=L/L_{\text{inel}}$) such subunits connected in series, which after ensemble averaging each contribute $N^{-1/2}$ (e^2/h) to the conductance and $N^{-3/2}$ (e^2/h) to the absolute conductance. Similarly, a 2D system can be thought of as being composed of M such quasi-1D channels connected in parallel. After averaging the relative conductance of each strip is $M^{-1/2}$ and the absolute contribution is $M^{1/2}$, so that the total conductance fluctuation for the entire system is $N^{-3/2}M^{1/2}$ (e^2/h). For Skocpol's particular device it was estimated that $N=7.5$ and $M=1.5$ giving a conductance fluctuation of about 0.04 (e^2/h) which was in reasonable agreement with the experimental result of 0.06 (e^2/h).

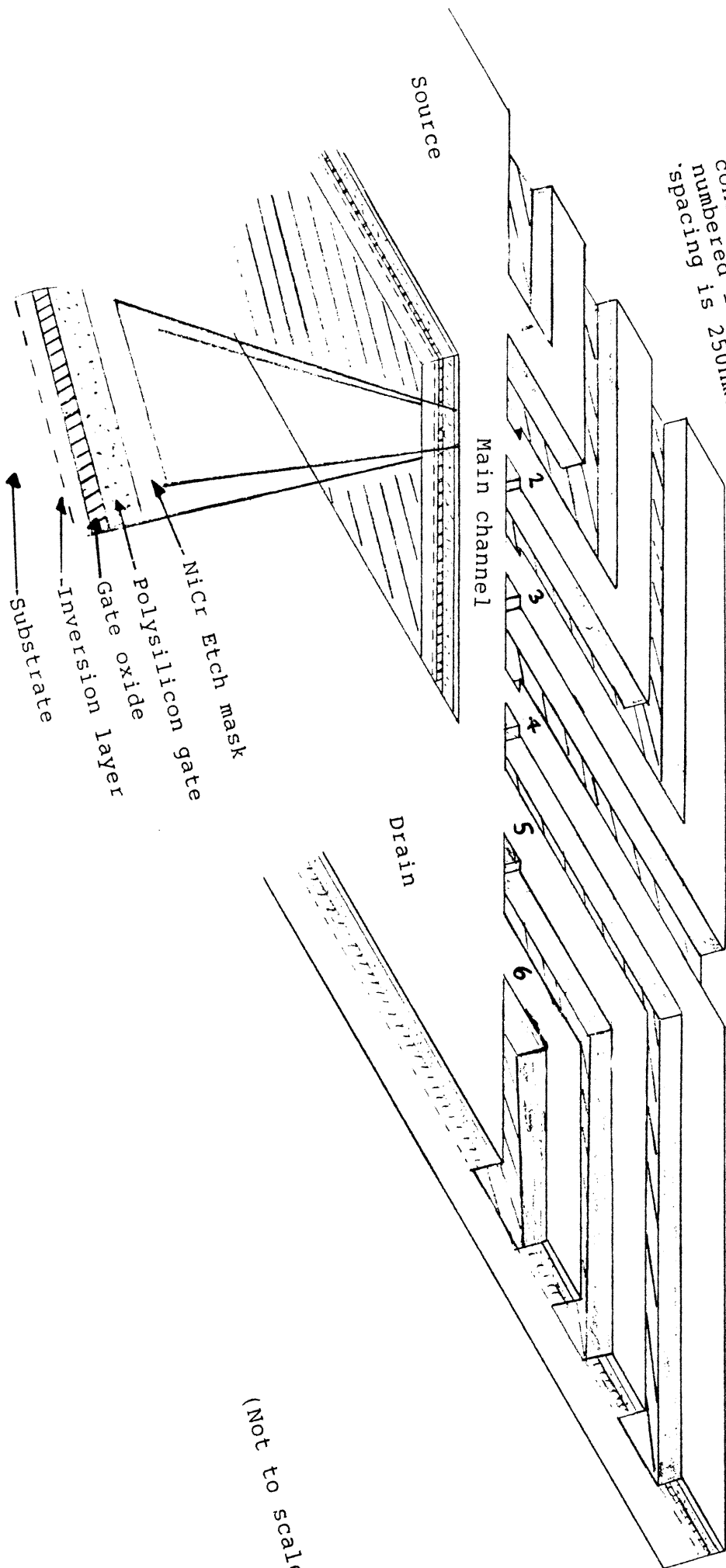
4.2.3 Electrostatic confinement of the 2DEG in an NMOS structure.

Another confinement approach was taken by Hartstein [1986] who used a MOSFET structure to study the physics of quasi-1D systems. Hartstein's structure, shown in figure 4.3, consists of a conventional FET of about 10 microns channel length with two additional control electrodes implanted either side. A negative voltage applied to the control electrodes enlarges the depletion layer around the p+ diffusions and thus acts to confine the channel to narrower widths allowing the width of the conducting channel to be varied between 20nm and 2 microns.

Using this device, a study of the channel conductance as a function of gate voltage for temperatures of 1.417K, 0.453K and 0.100K was performed (figure 4.4). Sharp fluctuations in the conductance were observed at the lowest temperature, which at first sight appear to be random noise, but were in fact found to be reproducible in a given sample (although they did vary from sample to sample). The amplitudes of the peaks in the conductivity were observed to increase with decreasing gate voltage or temperature.

The very sharp conductance oscillations in the low gate voltage regime were attributed to strong localisation arising from the statistics of just 1 or 2 hopping events whilst the oscillation in the higher gate voltage and higher temperature regime were attributed to weak localisation effects. A smooth transition between the two regimes was observed.

Figure 4.1
Schematic diagram of Skocpol's multi-terminal MOSFET. The main source-drain channel and the probe conducting channels numbered 1 to 6. The width of the main channel is 125nm and the probe spacing is 250nm.



(Not to scale)

Figure 4.2

The resistance of segment B of a multi-terminal MOSFET showing sharp changes in resistance as a result of the (thermal) activation of a single interface trap to act as a scatterer.

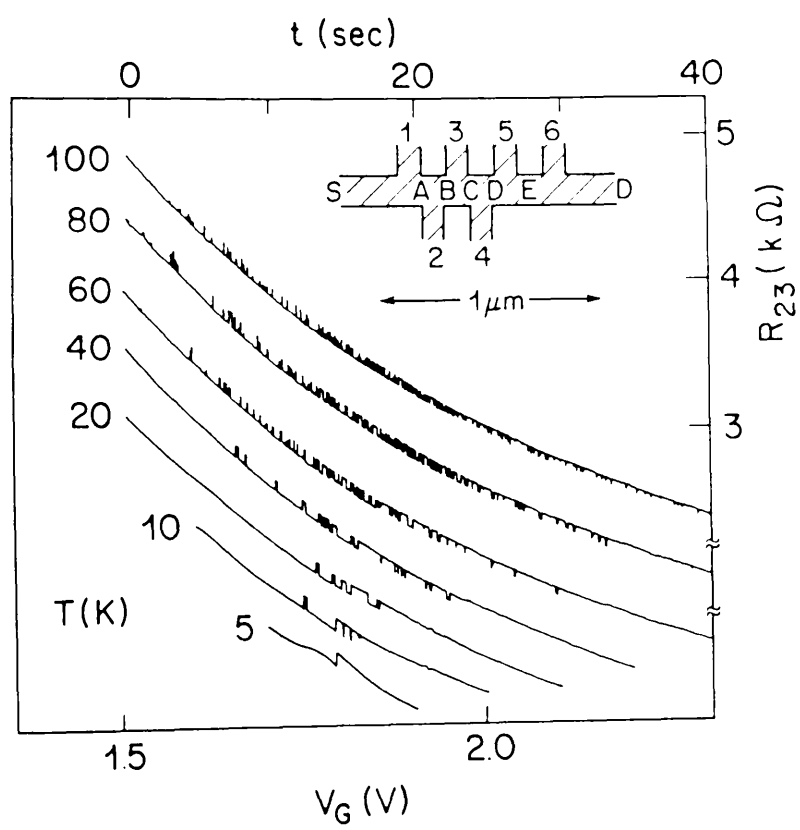


Figure 4.3

Schematic diagram of Hartstein's squeezed-channel FET.

(a) Plan view. The P⁺ regions are the control electrodes used to reduce the width of the channel and are about 1-2 microns apart and 14 microns long.

(b) Cross-sectional view along AA showing how the depletion regions confine the electrons in the inversion layer.

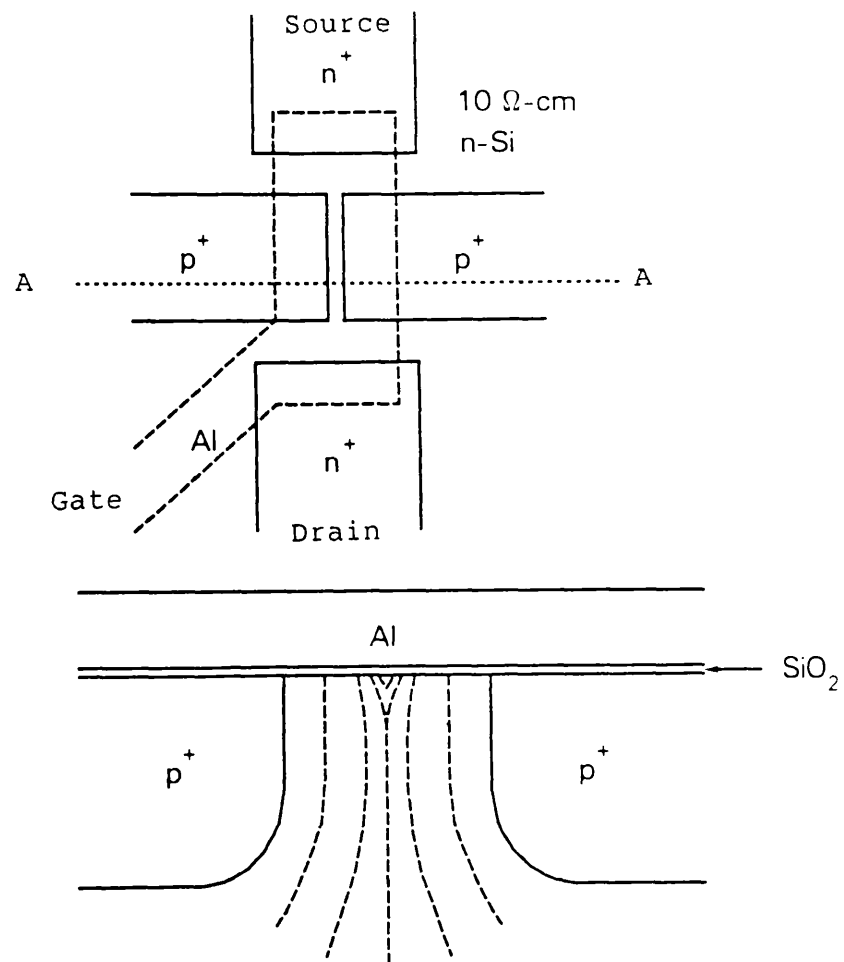
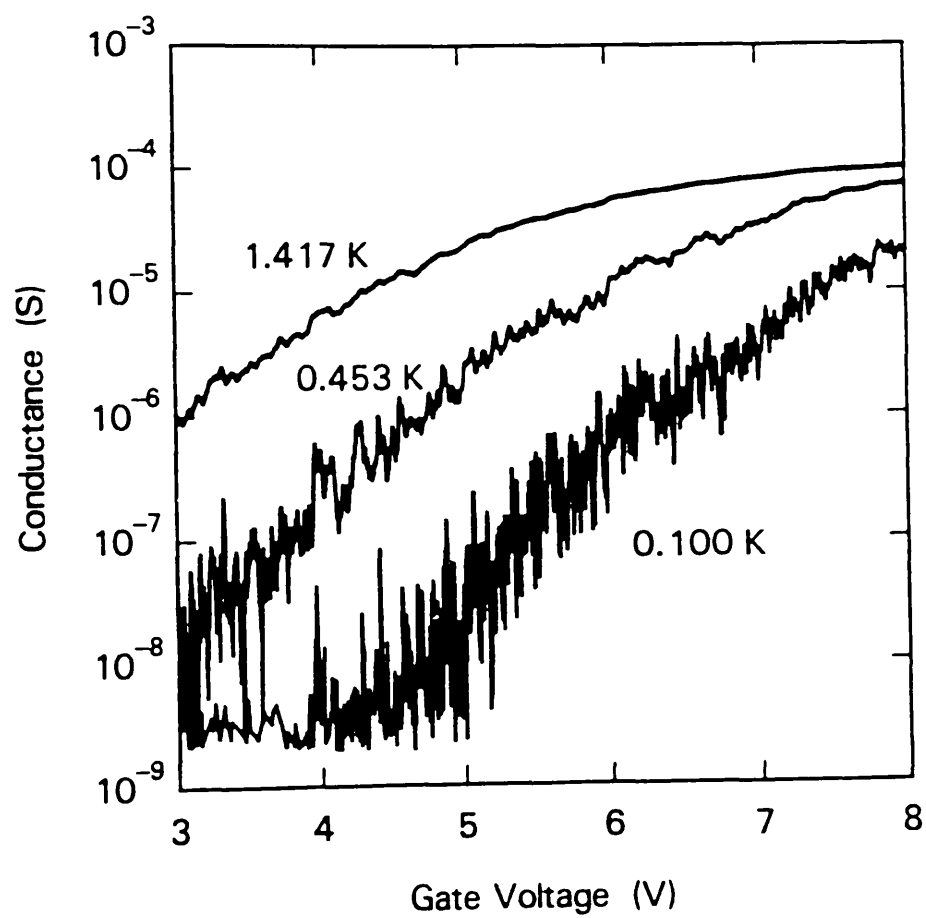


Figure 4.4

Typical conductance data from Hartstein's device showing large Conductance fluctuations as a function of gate voltage.



Comparison of the oscillations caused by the weak localisation was compared to the theory of Lee and Stone [1985] and was again found to be in good agreement, although it was commented that the degree of agreement was perhaps better than deserved considering the nature of the approximations made.

4.3 GaAs HEMT structures.

4.3.1 General

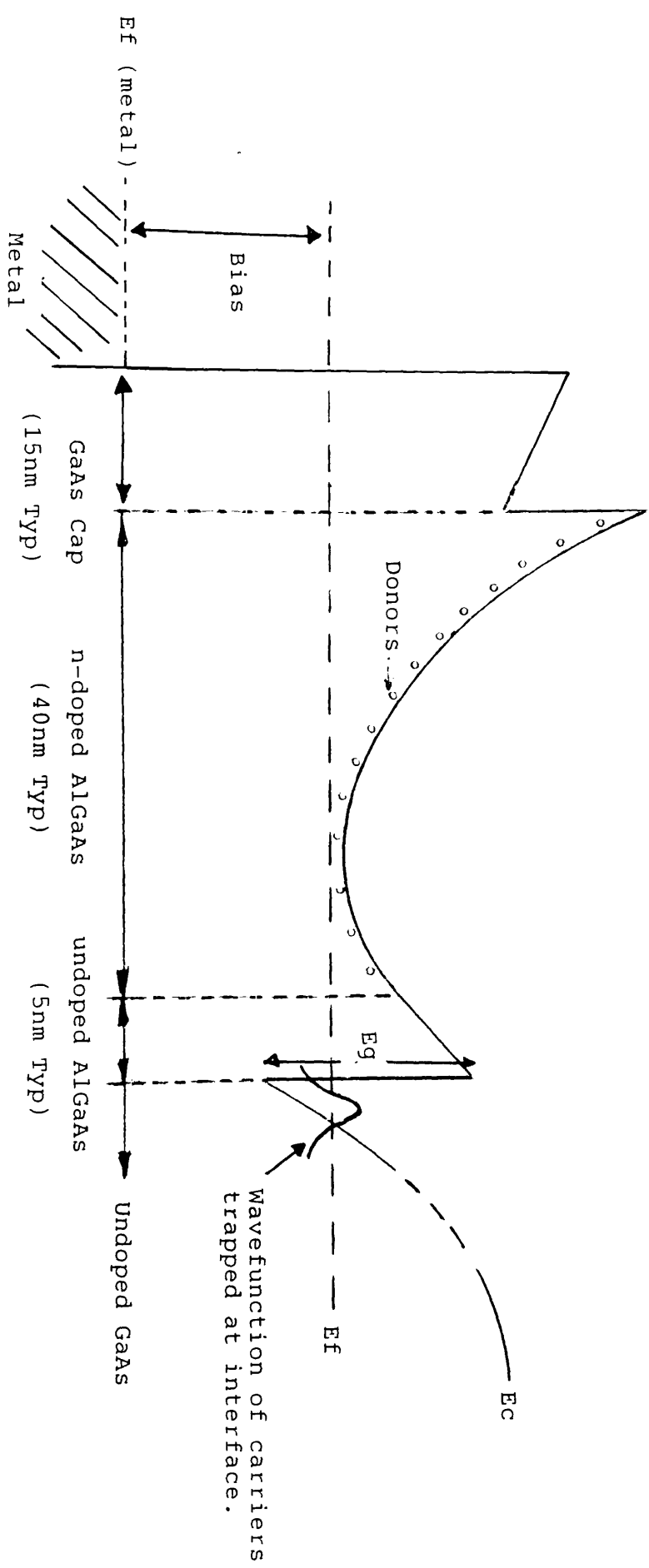
Another popular system for studying the physics of electron transport in one and two dimensions is in the two-dimensional electron gas (2DEG) present at a AlGaAs/GaAs interface within a HEMT structure. In this thesis, it is the electron transport through this type of structure which is of principal interest. In this section therefore the general principles of HEMT operation will be briefly reviewed for completeness and the use of these structures in the fabrication of AB rings will be described.

The high electron mobility transistor (HEMT) in its basic form consists of a highly doped AlGaAs layer, typically around 60nm thick on top of a nominally undoped GaAs layer. Source and drain contacts are then made, ensuring that electrical contact has been made to the active interface region, and a gate contact is then fabricated on the top surface.

4.3.2 Operational principles.

Operation of the HEMT is best understood by considering the conduction band profile vertically through the device. Figure 4.5 shows the conduction band profile after equilibrium has been attained when there are free carriers present at the heterojunction interface. The potential discontinuity at the semiconductor surface arises from the Schottky-barrier potential between the AlGaAs and metal when an electrode is present at the surface. When the surface is not contacted by an electrode but free, a potential discontinuity still occurs because the conduction band is held, or "pinned", at a fixed energy above the Fermi level in the bulk semiconductor due to the existence of surface states.

Figure 4.5 Conduction-band profile of a HEMT.



(Not to scale)

If the Fermi levels are considered to be flat and within the bulk of the metal and the semiconductor, then the application of a negative gate bias can be envisaged as an upward movement of the conduction band at the Schottky contact. This in turn, raises the rest of the conduction band with respect to the Fermi energy. When the gate is sufficiently negative with respect to the semiconductor, all of the conduction band is above the Fermi level, which means that there are no free carriers present and that the device is in its "off" state. In the "off" state, there are no surplus charges in the undoped GaAs and hence the conduction band in this region is flat. Decreasing the negative bias on the gate (*ie* making it more positive) brings the conduction band into closer proximity to the Fermi level and begins to mobilise carriers in the conduction band. Any free carriers mobilised in the doped region can then readily diffuse into the undoped GaAs, but not the other way around because of the potential discontinuity at the heterojunction.

When the system has reached equilibrium the familiar conduction band profile shown in figure 4.5 is established. Under these conditions, a high percentage of electrons transfer into the GaAs and are confined by the band-gap discontinuity and their own electrostatic potential into a thin sheet of charge, typically 10-100 nm thick, just underneath the heterojunction.

The roughly triangular shaped potential well in which the electrons find themselves, causes a quantisation of energy levels in the direction perpendicular to the layers (z-direction). Within each sub-band, transport is two-dimensional, with no motion being allowed in the z-direction and it is for this reason that the resulting sheet of charge is called a two-dimensional electron gas (2DEG).

The advantage of this arrangement is that the conduction electrons are now remote from their donor atoms, and therefore suffer less ionised impurity scattering. Indeed, they now move in a nominally undoped material which leads to a much higher mobility. However, since the 2DEG is still very close to the heterointerface, the remote donors can still cause some scattering which degrades performance. To reduce this effect, undoped spacer layers of AlGaAs are often grown between the highly doped region and the active layer in order to further separate the conduction electrons from their donors. If this technique is carried too far however, the spacer layer will start to inhibit charge transfer into the undoped GaAs and therefore a balance must be struck between high mobility but low carrier concentrations and low mobility and high carrier concentrations.

These issues have been addressed [Al-Mudares 1984] using a Monte Carlo model to find the optimum spacer layer thickness for a particular device. For the purposes of observing quantum effects, it is probably better to design the device for high mobility because even elastic scattering has the tendency to scramble the quantum effects (whilst not actually destroying them).

4.3.3 Confinement of the 2DEG in HEMT structures.

The uses of the HEMT structure however, need not be limited solely to normal transistor applications, and the major interest in this thesis derives from the fact that the 2DEG can be further confined into quasi-1D wires and ring structures.

Starting from the basic HEMT structure, several techniques have been employed to achieve the further confinement of the 2DEG and figures 4.6 to 4.9 show a schematic representation of the techniques which have been used to date.

The first method involves the etching of the semiconductor directly through the active interface thus forming rigidly defined wires (figure 4.6). There is however a variant of this which has been successfully used by Van Houten *et al* [1986] and others [Timp *et al* 1987]. This was to partially etch the top surface so that the etched surface was just above the heterojunction interface (figure 4.7). The surface potential arising from the pinning of the Fermi-level was then sufficient to deplete carriers from underneath the etched regions, and cause a confinement of carriers in the unetched regions.

A second very successful technique, which involves no etching, is to use a patterned gate contact to deplete the carriers in selected regions thus forming a shaped 2DEG (figure 4.8). The final method (figure 4.9) uses ion bombardment to damage selected areas of the material which substantially reduces the conductivity in those regions. Alternatively, the same technique can be used to locally increase the bandgap, causing the electrons to reside in the non-bombarded regions. This latter alternative however causes appreciable damage around the active interface region.

Generally, the shallow etch and gated-HEMT techniques of confinement have been favoured because of the minimal interference with the active interface region. Deep-etched wires however have been routinely fabricated [Leadbeater *et al* 1987] and the technique has also been used to produce AB ring structures [Ishibashi *et al* 1987].

The gated-HEMT structures have an additional advantage that the width of the conducting regions can be dynamically changed by the application of a gate voltage. The main disadvantage of the technique is that the gap between gates at the surface is large in relation to the "wire" widths. The shallow etch method does not allow the width of the wires to be changed once made, but could enable smaller structures to be manufactured [Ian McIntyre 1988].

The quantum-mechanical modelling of electron transport in these structures in this project was based on gated-HEMT structures, although the techniques developed could be easily applied to other structures.

The specific device modelled was an AB ring of the gated-HEMT type fabricated by Ford *et al* [1987]. The work of Ford *et al* was a development of earlier work by Berggren *et al* [1986] and Thornton *et al* [1986] who used a split-gate HEMT to study transport effects in quasi-1D channels.

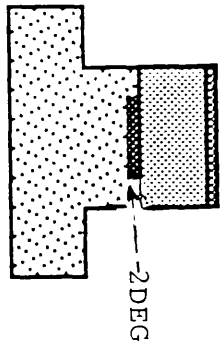
Figures 4.10 and 4.11 show a schematic diagram of an early ring fabricated by Ford *et al* [1987]. The conducting channels in this device could be made as narrow as about 200nm and possibly smaller and the AB oscillation amplitude was about 5% of the total resistance. The computer simulations by Finch [1987] were in general agreement with this figure and also suggested that the relatively wide exit to the ring would not favour very large AB oscillations due to the formation of many transverse modes (chapter 8). Therefore in a subsequent experiment, the exit from the ring was reduced from a width (as defined lithographically on the top surface) of about 1.3 microns to about 0.5 micron and the depth of the 2DEG from the top surface reduced from 150 to 60nm (figure 4.12). The amplitude of the oscillations in the latter ring were as large as 18% of the background resistance, suggesting that the number of transverse modes (which, it is proposed, lead to the loss of modulation depth) had been successfully reduced.

These experimental results and the associated computer simulations will be discussed in greater detail in chapters 8 and 9.

Figures 4.6 to 4.9

Fabrication techniques used to achieve confinement of the 2DEG in HEMT-structures.

Figure 4.6
Deep etched



Etched structures

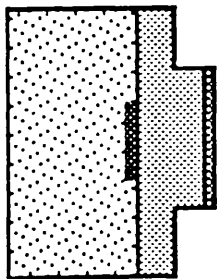
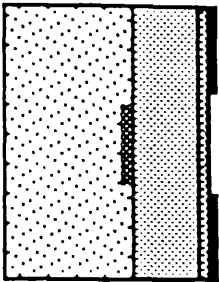


Figure 4.7
Shallow etched

Figure 4.8
Patterned gate



GaAs
AlGaAs

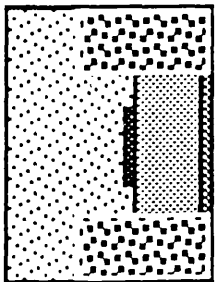
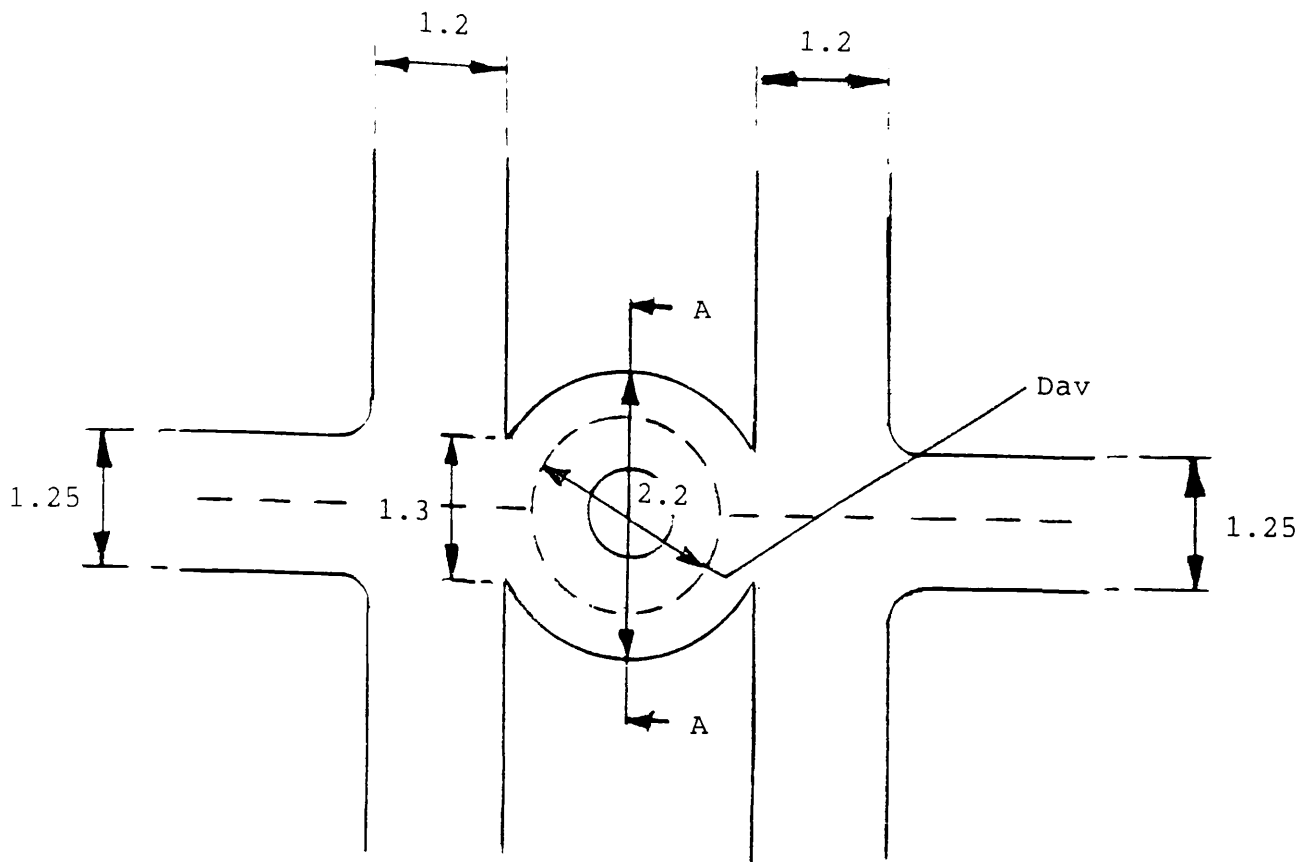


Figure 4.9
Implanted or damaged

GaAs/substrate
Implanted/damaged region



D_{av} = Average diameter of conducting ring ≈ 1.3 micron.

The width of the channel in the 2DEG is about 200nm

Average diameter of central hole ≈ 0.37 micron

Figure 4.10

Schematic diagram of a patterned-gate AB device fabricated by Ford et al as it appears on the top surface (ie dimensions of channels are lithographic and not actual conducting widths).

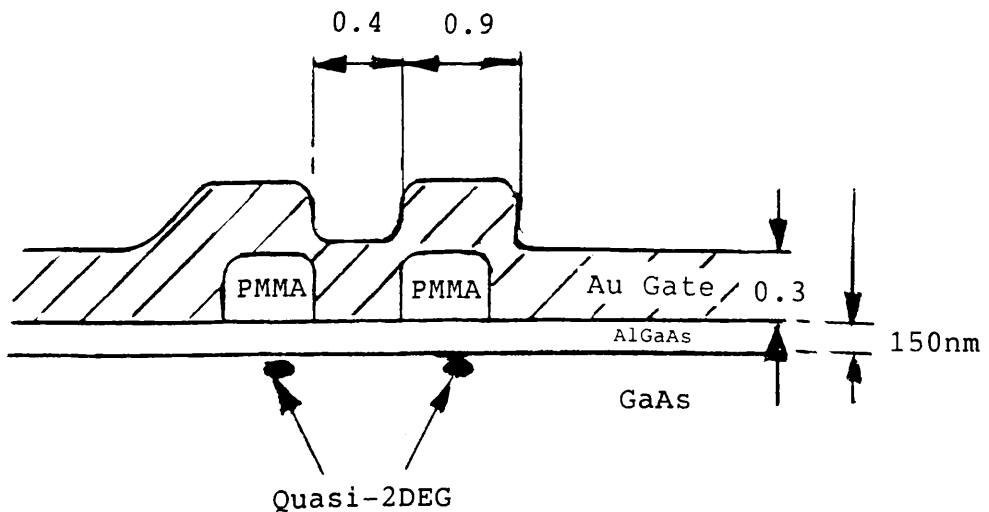
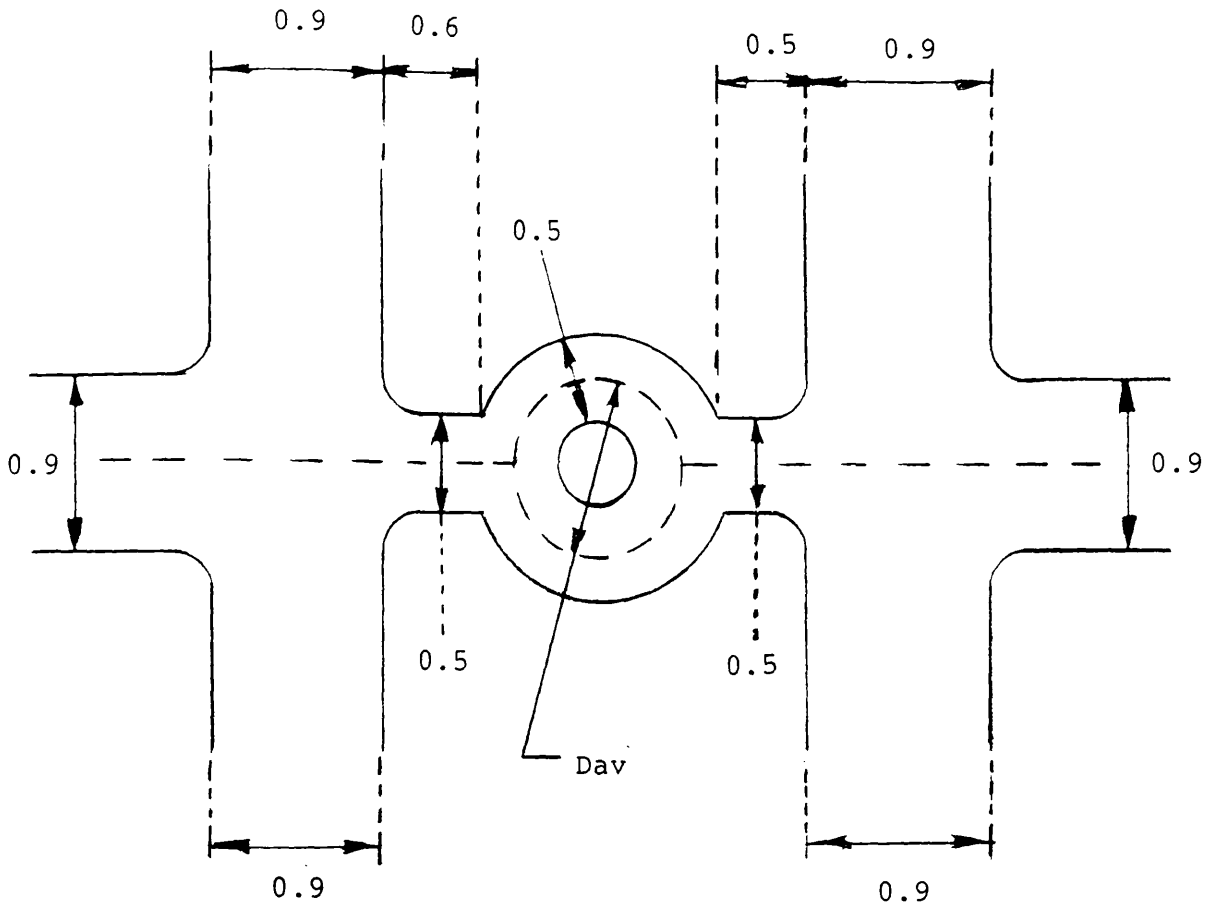


Figure 4.11 Cross-sectional view along AA for the above device.



Average diameter of central hole \approx 0.86 micron

D_{av} = Average diameter of conducting ring \approx 1.36 micron

The quasi-2DEG channel is about 60nm below the surface
and is about 100nm wide.

Figure 4.12

Improved ring structure fabricated by Ford et al.
This ring structure produced magneto-resistance
oscillation amplitudes of up to 20% of the background resistance.

4.4 Summary.

This chapter has briefly described how semiconducting systems for confining the conduction electrons in one or two dimensions can be realised. The main motivation for the use of HEMT structures is that the high mobility in the 2DEG makes these structures good candidates for the observation of coherent quantum interference effects such as the AB effect.

Therefore particular attention was paid to methods of confining the 2DEG in HEMTs. From the viewpoint of AB ring fabrication, the shallow etch and gated-HEMT techniques seem to be the best methods as both involve minimal damage to the active interface region.

Ultimately the shallow etch technique could produce the best results as it appears that smaller rings could be manufactured this way. At the time this work was in progress however, the shallow-etch technique had not yet been perfected and the best available results were being obtained with gated-HEMT structures, so modelling proceeded with this system.

5. Modelling of an Aharonov-Bohm ring with one-dimensional conductors.

5.1 Introduction.

This section shows how the technique for calculating transmission coefficients using transmission matrices can be extended to include multiply-connected one-dimensional structures, such as rings.

Although these models may be far removed from experimental reality, they do provide a means of understanding much of the basic behaviour of such structures at a simple level. These models then, can be of use in the interpretation of experimental data provided that the validity and limitations of the model are kept in mind.

An explicit formula for the transmission coefficient of an AB ring is developed using transmission and scattering matrices to match the plane wave coefficients around the ring.

The transmission coefficient of the ring is studied as a function of enclosed magnetic flux and wavevector for different model parameters.

The effect of elastic scattering in the arms of the ring is studied qualitatively and the large variations in the behaviour of the ring is demonstrated.

5.2 Elements of the model.

The one-dimensional model for an AB ring was composed of two main elements;

- 1) The transmission matrix for a 1D potential.
- 2) The scattering matrix for a 3-port Y-junction.

The transmission matrix (T-matrix) expresses the forward and backward propagating plane-waves on the right hand side of the potential barrier in terms of those on the left hand side. A detailed description of the transmission matrix was given in chapter 2 (and appendix A). As a corollary, it is seen that the T-matrix is able to describe how a plane wave will propagate through any potential distribution, so that the free-space case is also included.

The scattering matrix (S-matrix) is similar to the transmission matrix in that it deals with plane waves at a junction, but dissimilar in that it relates the outgoing waves from a junction to the incoming ones.

The form of the S-matrix is derived by assuming that the matrix must be unitary and exhibits time-reversal symmetry. The unitarity of the S-matrix ensures that current is conserved in the scattering process and hence the sum of all incoming currents must equal the sum of all outgoing currents.

The second assumption of time-reversal symmetry is a little more restrictive. The time-dependent Schrodinger equation for particles moving in a magnetic field does not satisfy the test for time-reversal invariance [Landau and Lifshitz 1958], namely that of testing if the Schrodinger equation retains the same form if the sign of 't' is reversed and at the same time the complex conjugate of the equation is taken. The test for time-reversal invariance fails because the Hamiltonian for a particle moving in a magnetic field has an imaginary part. However, for the case in which the magnetic field is confined solely to the interior of the ring, it will be shown that the Schrodinger equation can be reduced to that for a free particle if the wavefunction is multiplied by an additional phase factor arising from the magnetic vector potential. Since the free-particle Hamiltonian is real, the Schrodinger equation for the ring with an enclosed flux does exhibit time-reversal invariance and therefore all the results in this section are concerned with this situation.

The approach taken by Buttiker *et al* was to further simplify the S-matrix by considering all of the elements to be real; a procedure that was justified by the arbitrary division of elastic scattering and free propagation between the S and T matrices. All phase factors pertaining to propagation and scattering in the arms of the ring were therefore represented by the T-matrices. This means that the y-junction may be thought of as a singular point off which the incident wavefunction scatters.

Alternatively, one could use the approach of Gefen *et al* [1984] which was to incorporate scattering and propagation terms for the arms of the ring into the definition of the S-matrix thus eliminating the T-matrices from the problem. Although this model for the ring required only two S-matrices joined back to back, the form of the matrices were considerably more complex so it was doubtful whether any real simplification had been made.

In this project it was decided to use Buttiker's strategy as it seemed more modular and hence extendable to more complex networks other than rings. The modular nature of the procedure also made it suitable for computer implementation.

Enforcing unitarity on the S-matrix together with the requirement that it must also be symmetric, leads to the following general relationships:-

$$S_{11}^2 + S_{12}^2 + S_{13}^2 = 1 \quad (\text{Eqn 5.1})$$

$$S_{21}^2 + S_{22}^2 + S_{23}^2 = 1 \quad (\text{Eqn 5.2})$$

$$S_{31}^2 + S_{32}^2 + S_{33}^2 = 1 \quad (\text{Eqn 5.3})$$

$$S_{11}S_{21} + S_{12}S_{22} + S_{13}S_{33} = 0 \quad (\text{Eqn 5.4})$$

$$S_{11}S_{31} + S_{12}S_{32} + S_{13}S_{33} = 0 \quad (\text{Eqn 5.5})$$

$$S_{21}S_{11} + S_{22}S_{12} + S_{23}S_{13} = 0 \quad (\text{Eqn 5.6})$$

$$S_{21} = S_{12} \quad (\text{Eqn 5.7})$$

$$S_{31} = S_{13} \quad (\text{Eqn 5.8})$$

$$S_{32} = S_{23} \quad (\text{Eqn 5.9})$$

To make any further progress more information about the S-matrix is needed. If the matrix element S_{ij} is interpreted as the coupling between wire i and wire j , then taking wire 1 to be the "input" wire, S_{12} and S_{13} represent the coupling coefficients between the input wire and each of the arms forming the ring. If the ring is considered to be symmetrical with respect to the input wire, then there is now enough information available to assemble the complete S-matrix. The coupling between the input and the arms, denoted by e , can vary between 0 and 1/2 representing the case when the ring is poorly and strongly coupled to the external leads respectively.

Since however, the S-matrix in this model deals with the wavefunctions and not the probability densities, S_{12} and S_{13} are equated with $\epsilon^{1/2}$.

Inserting these results into the relationships expressed by equations 5.1 to 5.9 yields an S-matrix of the form:-

$$\mathbf{S} = \begin{pmatrix} -(a+b) \epsilon^{1/2} & \epsilon^{1/2} \\ \epsilon^{1/2} & a & b \\ \epsilon^{1/2} & b & a \end{pmatrix} \quad (\text{Eqn 5.10})$$

A form obtained by Buttiker *et al* [1984].

Solving for 'a' and 'b' between equations 5.1 to 5.9 and 5.10 yields the result:-

$$a_{+/-} = (+/-) \frac{1}{2} \left((1-2\epsilon)^{1/2} - 1 \right) \quad (\text{Eqn 5.11})$$

$$b_{+/-} = (+/-) \frac{1}{2} \left((1-2\epsilon)^{1/2} + 1 \right) \quad (\text{Eqn 5.12})$$

5.3 The transmission coefficient of a ring with one-dimensional conductors.

The ring to be modelled is considered to be constructed of two S-matrices and two T-matrices connected as shown in figure 5.1. As for the case of the T-matrix for the 1D potential (Appendix A), a solution can only be obtained by inserting an appropriate "boundary condition", in this case $f=0$ and $a=1$. This condition represents a plane wave of unity amplitude incident on the ring from a point far to the left of the ring. The coefficient f' is set equal to zero as there will be no back-reflection in this idealised model since the output wire is considered to be infinite in extent. If necessary a boundary on the right hand side of the ring could be modelled, but this would require the inclusion of another T-matrix.

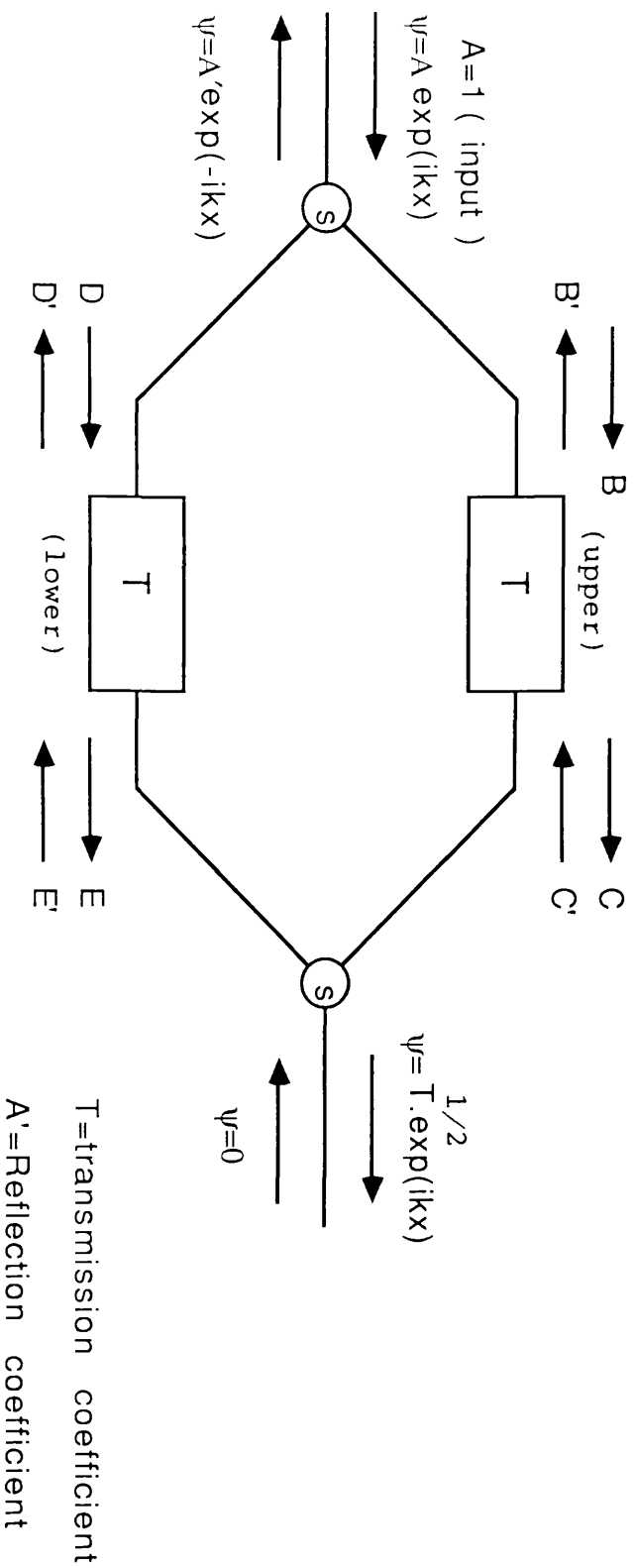


Figure 5.1 S&T Matrix method used to calculate the transmission of a ring with one-dimensional conductors.

The arrows represent the direction of propagation of the plane waves that must be matched at each interface. The letters represent the coefficients of the plane waves in each region.

Buttiker's and Imry's calculation involved the (legitimate) swapping around of T-matrix elements to simplify the calculation. In this project it was intended to incorporate the previous work on T-matrices into the model and also use standard definitions so that a modular structure could be set-up for more complex networks. Whilst the use of non-standard definitions did not present an unsurmountable problem, the author can see no reason why the transmission could not be calculated directly in terms of the standard definitions outlined in chapter 2. The following pages outline the various stages of the calculation.

The procedure is to assume that a plane wave of unit amplitude is incident from the left, and with the aid of the S and T-matrices, calculate the amplitude of the wave emerging at the exit. Since the waves are assumed to be incident from the left only, the coefficient for the backward-propagating wave at the exit of the ring is set equal to zero. The system of equations to be solved is therefore as follows:-

S-matrix equation for the input y-junction.

$$\begin{pmatrix} A' \\ B' \\ C' \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ B \\ C \end{pmatrix} \quad (\text{Eqn 5.13})$$

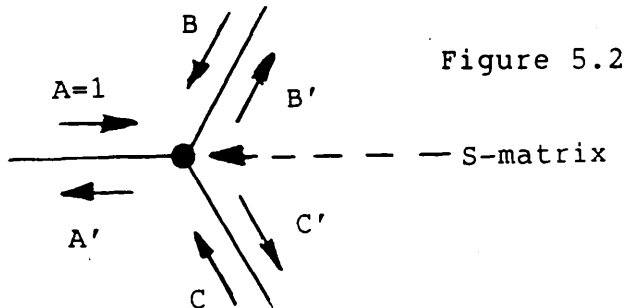


Figure 5.2

T-matrix equation for the "upper" arm.

$$\begin{pmatrix} D \\ D' \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} \begin{pmatrix} B' \\ B \end{pmatrix} \quad (\text{Eqn 5.14})$$

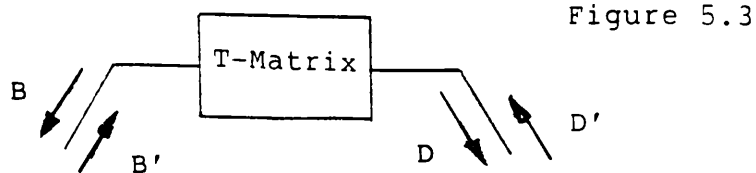


Figure 5.3

T-matrix equation for the "lower" arm.

$$\begin{pmatrix} E \\ E' \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} C' \\ C \end{pmatrix} \quad (\text{Eqn 5.15})$$

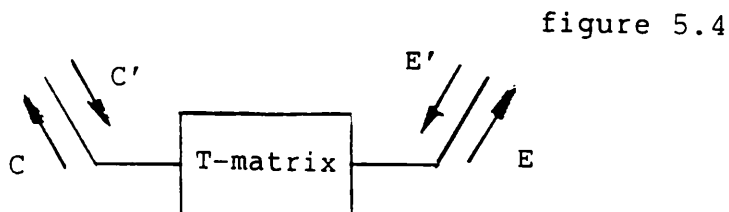


figure 5.4

S-matrix equation for the y-junction.

$$\begin{pmatrix} F' \\ E' \\ D' \end{pmatrix} = \begin{pmatrix} S'_{11} & S'_{12} & S'_{13} \\ S'_{21} & S'_{22} & S'_{23} \\ S'_{31} & S'_{32} & S'_{33} \end{pmatrix} \begin{pmatrix} 0 \\ E \\ D \end{pmatrix} \quad (\text{Eqn 5.16})$$

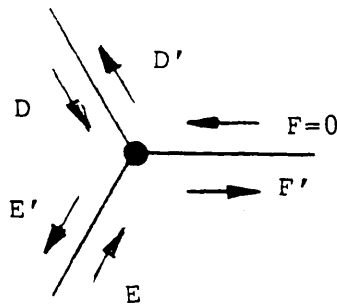


Figure 5.5

Although these equations can be solved for the general case, working may be simplified by assuming that both the S-matrices are identical and that (in the first instance) the T-matrices are those representing free-space propagation along the arms of the ring.

Solving this system of equations, one obtains the following results for the transmission and reflection coefficients of the ring.

$$t = \frac{\begin{vmatrix} \alpha_1 & \gamma_1 \\ \alpha_3 & \gamma_3 \end{vmatrix} \begin{vmatrix} \beta_2 & \gamma_2 \\ \beta_3 & \gamma_3 \end{vmatrix} + \begin{vmatrix} \alpha_2 & \gamma_2 \\ \alpha_3 & \gamma_3 \end{vmatrix} \begin{vmatrix} \beta_3 & \gamma_3 \\ \beta_1 & \gamma_1 \end{vmatrix}}{\gamma_3 \begin{vmatrix} \beta_2 & \gamma_2 \\ \beta_3 & \gamma_3 \end{vmatrix}}$$

(Eqn 5.17)

$$r = \frac{\begin{vmatrix} S_{11} & \gamma_1 \\ S_{13} & \gamma_3 \end{vmatrix} \begin{vmatrix} \beta_2 & \gamma_2 \\ \beta_3 & \gamma_3 \end{vmatrix} + \begin{vmatrix} S_{12} & \gamma_2 \\ S_{13} & \gamma_3 \end{vmatrix} \begin{vmatrix} \beta_3 & \gamma_3 \\ \beta_1 & \gamma_1 \end{vmatrix}}{\gamma_3 \begin{vmatrix} \beta_2 & \gamma_2 \\ \beta_3 & \gamma_3 \end{vmatrix}} \quad (\text{Eqn 5.18})$$

Where α_n , β_n and γ_n are given by:-

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = (S_{12}, S_{13}) \begin{pmatrix} T_{11} & 0 \\ 0 & t_{11} \end{pmatrix} \begin{pmatrix} S_{31} & S_{32} & S_{33} \\ S_{21} & S_{22} & S_{23} \end{pmatrix} \quad (\text{Eqn 5.19})$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} = (S_{22}, S_{23}) \begin{pmatrix} T_{11} & 0 \\ 0 & t_{11} \end{pmatrix} \begin{pmatrix} S_{31} & S_{32} & S_{33} \\ S_{21} & S_{22} & S_{23} \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ T_{22} \end{pmatrix} \quad (\text{Eqn 5.20})$$

$$\begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} = (S_{32}, S_{33}) \begin{pmatrix} T_{11} & 0 \\ 0 & t_{11} \end{pmatrix} \begin{pmatrix} S_{31} & S_{32} & S_{33} \\ S_{21} & S_{22} & S_{23} \end{pmatrix} - \begin{pmatrix} 0 \\ t_{22} \\ 0 \end{pmatrix} \quad (\text{Eqn 5.21})$$

where the elements t_{nm} belong to the transmission matrix for the upper arm of the ring and T_{nm} to the lower arm.

The "upper" T-matrix for free propagation along a one-dimensional wire is of the form:-

$$t(k, \phi) = \begin{pmatrix} \exp(i(kl + \phi)) & 0 \\ 0 & \exp(-i(kl + \phi)) \end{pmatrix} \quad (\text{Eqn 5.22})$$

The length, l , is the distance around the upper half of the ring from the input to output junction and ϕ is the flux-induced phase change. The form of the "lower" T-matrix is similar, but the sign of ϕ is reversed as the wavetrain is propagating around the region of flux in the opposite sense to the upper arm. The "lower" T-matrix thus has the form:-

$$T(k, \phi) = \begin{pmatrix} \exp(i(kl - \phi)) & 0 \\ 0 & \exp(-i(kl - \phi)) \end{pmatrix} \quad (\text{Eqn 5.23})$$

For a ring without scatterers in the arms and a coupling parameter of 0.5, equation 5.17 reduces to:-

$$t(k, \phi) = (e^{2ikl} (e^{2ikl} - \cos^2 \phi) - \sin^2 \phi) / (e^{ikl} \cos \phi (e^{2ikl} - \cos^2 \phi))$$

$$\text{with } \phi = \pi(\Phi / (h/e)). \quad (\text{Eqn 5.24})$$

Even for a perfectly coupled ring without any scatterers, the transmission characteristic is no longer a simple sinusoidal function of the enclosed flux which was the case in the two-slit calculation of chapter 3. The other major difference is that there is now a functional dependence on the wavevector through the appearance of terms involving the product kl . These terms give rise to energy resonances, which were also not present in the simple two-slit calculation.

5.4 Results.

The derived transmission formula (equation 5.17) was implemented on an HP9000 desktop computer and the behaviour of the transmission as a function of flux, coupling parameter and energy was studied. In this description of the ring, it is seen that the actual values of the wavevector, k , and half-circumference, l , are unimportant separately, and it is only their product which is of significance. The transmission coefficient was therefore plotted as a function of the ' kl ' product. As the transmission function will be periodic in kl , only kl values between 0 and 2π were considered.

The behaviour of the transmission as a function of energy with fixed flux values and coupling parameters was considered first. Two distinct regimes become apparent. The well-coupled regime (with the coupling parameter close to a half) exhibits a transmission characteristic that is near unity most of the time for small values of applied flux. For energy values where the kl product is close to an integer multiple of π , the transmission falls sharply to zero. As the applied flux is increased, the zero-points in the transmission profile persist, but the peaks become lower and smoother (figures 5.6 to 5.9). These eventually tend towards zero for all wavevectors as the applied flux tends towards a value of $1/2(h/e)$ and shows the occurrence of the AB effect in one dimensional rings.

It is known that it is not possible to perfectly match each of the y -junctions individually, but it can be seen from the graphs that the coupling to the rest of the ring system means that the ring as a whole can be matched, resulting in unity transmission at certain values of wavevector.

As the coupling parameter ε decreases the peaks in transmission become more singular (figures 5.10 and 5.11), and in the limit of ε tending to zero, these energies may be identified as the eigenstates of the isolated ring. For coupling parameters less than about 0.1 it was found that the energies and shifts in the energies with applied flux agreed well with a model for an orbiting electron in a magnetic field [Olariu and Popescu 1985].

5.5 Discussion of results: A phasor diagram approach.

The positions of the peaks and their amplitudes in both the high and low coupling regime can be understood in terms of the freely orbiting electron model [Olariu and Popescu 1985] or from the work of Buttiker *et al* [1984]. What was not so clear was why the transmission persistently became zero at regular intervals of $kl=n\pi$. In constructing a reasonable theory for this, it is important to notice that the transmission is always exactly zero (to within machine accuracy) at these values, and is not simply a small value. Also, these values always occur when $kl=n\pi$ and are flux independent, which would seem to rule-out any multiple-reflection process.

The second study performed was to investigate the behaviour of the transmission as a function of the enclosed flux. Figures 5.12 to 5.14 show the transmission coefficient of a perfectly coupled ring as a function of ϕ/ϕ_0 , where ϕ_0 is the h/e flux quantum. For coupling parameters close to 0.5, the transmission curves are of approximately cosinusoidal form with minima at ϕ/ϕ_0 as expected and are rather insensitive to variations in the kl -product.

If the coupling parameter is now reduced, the behaviour of the transmission coefficient becomes more complex. In similarity with the study of the wavevector dependence, resonances in the transmission occur at certain flux values (Figures 5.15 and 5.16). In the case where $\varepsilon=0.25$ and $kl=0.35*(2\pi)$ for instance, the transmission is seen to initially rise, which is quite different to the initial reduction one would expect from the simple two-slit model which does not take energy resonances into account.

A better understanding of the behaviour of the ring in the well-coupled regime can be gained by representing the plane waves around the ring in vector form on a phasor diagram. Referring to figure 5.17 representing the strong coupling case, the two vectors of equal length represent the wavefunctions in the upper and lower arms.

It can be seen that for a ring with equal arm lengths, the two phasors (representing amplitude and phase of the wavefunction at the output) are symmetrically disposed about the real axis when the kl product around the arms is $n\pi$. The phases ϕ_+ and ϕ_- are the flux-induced phase shifts for the upper and lower arms (taking the zero-phase reference point to be at the input junction) and are equal and opposite because the paths they refer to are identical but encircle the flux in opposite senses. It is now seen that when the kl product is $n\pi$ the resultant wavevector always lies along the real axis

independent of the applied flux. Since the S and T matrices match the coefficients of plane waves around the ring, the transmission coefficient thus obtained is in terms of currents. This then, explains the sharp drops in transmission when $kl=n\pi$, because the resultant wavefunction at the output in these instances is purely real and hence cannot carry any current.

Considering the low coupling regime, there are now many phasors on the diagram (though of reduced length to maintain overall normalisation). These phasors are produced each time the wavetrain cycles around the ring, being split at each pass through the input and output S-matrices. A wavefunction internal to the ring is still split into transmitted and reflected portions even if the coupling parameter is 1/2 (representing perfect coupling). This is because the single S-matrix cannot be perfectly matched for plane waves both internal and external to the ring, as was remarked upon earlier.

Figure 5.18 shows the steady state diagram for a ring of low coupling parameter and arbitrary flux linkage. The smaller the coupling parameter, the shorter and more numerous the phasors. For most flux values, the phasors will have travelled different distances and had their phases changed by varying amounts according to the path taken. Thus on average, when the phasors are added together at the output junction, the resultant will be small. This is because the more phasors there are, the more symmetrical their distribution about the origin, and hence the resultant will also be near the origin. In the limit of an infinite number of phasors evenly distributed about the origin, the resultant will generally be zero. There are however exceptions, where the transmission can rise to unity. These situations occur when all the partial phasors add together coherently when the phase shift due to the flux ϕ and that due to the freely propagating wave (kl) add up so that:-

$$2kl + 2\phi = 2n\pi \quad (\text{Eqn 5.25})$$

This is so that on each cycle around the ring the total phase advancement of the next partial phasor is such that it is at the same position on the diagram as it was in the previous cycle. This ensures that all partial phasors add constructively.

Figure 5.16 shows a plot of transmission versus flux linkage for a low coupling parameter and clearly shows the peaks in transmission at the allowed kl -values. For a ring with a flux linkage of Φ , the relation

$$2\phi = (e/h)\Phi \quad (\text{Eqn 5.26})$$

may be substituted into equation 5.25 to obtain an expression for the positions of the energy resonances:-

$$kl/(2\pi) = 1/2 (n - \Phi/\Phi_0) \quad \text{with} \quad \Phi_0 = h/e \quad (\text{Eqn 5.27})$$

which is the same result that would have been obtained by applying the Sommerfeld quantisation condition to quantise the canonical angular momentum [Olariu and Popescu 1985].

In the more general situation, ϵ is rather less than 0.5 but still significant (figures 5.10 and 5.11). Since there are fewer phasors in this case, complete cancellation does not occur for all the non-allowed kl -values and the peaks are not as sharp. Zero transmission does however always occur on the AB condition when pairs of partial waves are out of phase with each other by 180 degrees. This condition is demonstrated by figure 5.9 which shows that the transmission is tending to zero for all kl values as the flux tends towards $0.5\Phi_0$. A numerical calculation of the transmission for a flux linkage of exactly $0.5\Phi_0$ was not possible due to a singularity at this value. It was however possible to get as close as machine accuracy would allow, so that the limit $\Phi \rightarrow 0.5\Phi_0$ could be approached.

The behaviour of the transmission as a function of flux for a coupling parameters of between 0.5 and 0.1 can be similarly understood. In particular, the initial rise in transmission for coupling parameters of about 0.2-0.35 can be attributed to the ring slowly approaching resonance. This type of behaviour is noted in later chapters which consider a two-dimensional model of the ring. The conclusion to be drawn from these findings is that the coupling parameter in the particular two-dimensional model considered is probably in the range 0.2-0.35.

A major difficulty with the 1D model is that the choice of coupling parameter is quite arbitrary within the range 0-0.5. An objective of later work using two-dimensional models therefore, was to obtain a less arbitrary estimate of the 1D model parameters by fitting the 1D transmission characteristics to the 2D magneto-transmission curves.

5.6 A Qualitative study of the effects of impurities in rings with one-dimensional conductors.

In this section the effects of elastic scattering in the arms of the ring are investigated. This is likely to be important in real rings as a result of the potentials arising from remote ionised donors.

Elastic scattering in the arms was modelled by the introduction of transmission matrices of the form [Buttiker *et al* 1984, Appendix(A)].

$$t = \begin{pmatrix} 1/t^* & -r^*/t^* \\ -r/t & 1/t \end{pmatrix} \quad (\text{Eqn 5.28})$$

$$\text{with } t = |T_S|^{1/2} \exp(i\phi_S) \quad (\text{Eqn 5.29})$$

where T_S is the square-modulus of the transmission coefficient of the scatterer and ϕ_S is the phase change in the transmitted wave. In this work it is only the effect of altering T_S which is investigated.

The results obtained agreed with the theory of Buttiker *et al* [1984] who related the peaks in the transmission coefficient to the poles in transmission formula. In the case of strong coupling to the external leads, the pole in the transmission coefficient converged to a flux value of $\phi/\phi_0 = 3\pi/2$ as the elastic scattering increased. What was not perhaps stressed sufficiently in the work of Buttiker *et al* was the large variation in behaviour when elastic scattering was incorporated into the model. The main purpose here therefore, is to indicate the variety of behaviour possible and to point out some general features of the transmission curves so that they can be used in the interpretation of the two-dimensional simulation results in chapter 9.

Figures 5.19 to 5.21 show the effect of altering the transmission of one of the arms of the ring. The curves can be seen to be progressively distorted from the case when the arm transmissions are unity. When both arm transmissions are unity the curves are always symmetric about $kl=0.5*(2\pi)$, but as one of the arm transmissions is reduced, this symmetry is destroyed although the curves are still periodic (with period 2π).

When both arm transmissions are changed (figures 5.22 to 5.24) the symmetry in the transmission curves is retained, but only one axis of symmetry remains at $kl=0.75*(2\pi)$. As the arm transmissions are reduced, the peaks in the ring transmission either side of the symmetry axis become progressively sharper and the transmission at other kl -values away from the resonance condition becomes lower. When there is no special relationship between the partial waves traversing the ring, they add as though they are incoherent and the net transmission is the same as the transmission of the individual matrices and is indifferent to the fact that the waves have travelled along different paths to reach the output.

Even the effects of quite modest scattering in the arms can have a drastic effect on the transmission characteristic. Figure 5.25 shows the effect of elastic scattering in a well-coupled ring with an enclosed flux of $\phi/\phi_0=0.35$. This graph is greatly distorted about the value $kl=0.5*(2\pi)$ compared to the case without any scattering (figure 5.7). Figures 5.25 to 5.27 show the effect of altering the coupling parameter for a fixed arm transmission. As the coupling parameter is decreased, some asymmetry is still observed in the plots, but the main effect is to expand the distance between the first two peaks and reduce the distance between the second pair of peaks.

Finally, figures 5.28 and 5.29 show the transmission coefficients of a perfectly coupled ring with a kl product of $0.1*2\pi$. These plots demonstrate that in certain circumstances the effect of the scattering can be to suppress the zero in transmission altogether. Indeed, the transmission can actually peak at a flux value of around $0.5*(h/e)$; a complete reversal of the behaviour of the ring without any scatterers.

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .1 * \Phi_0$

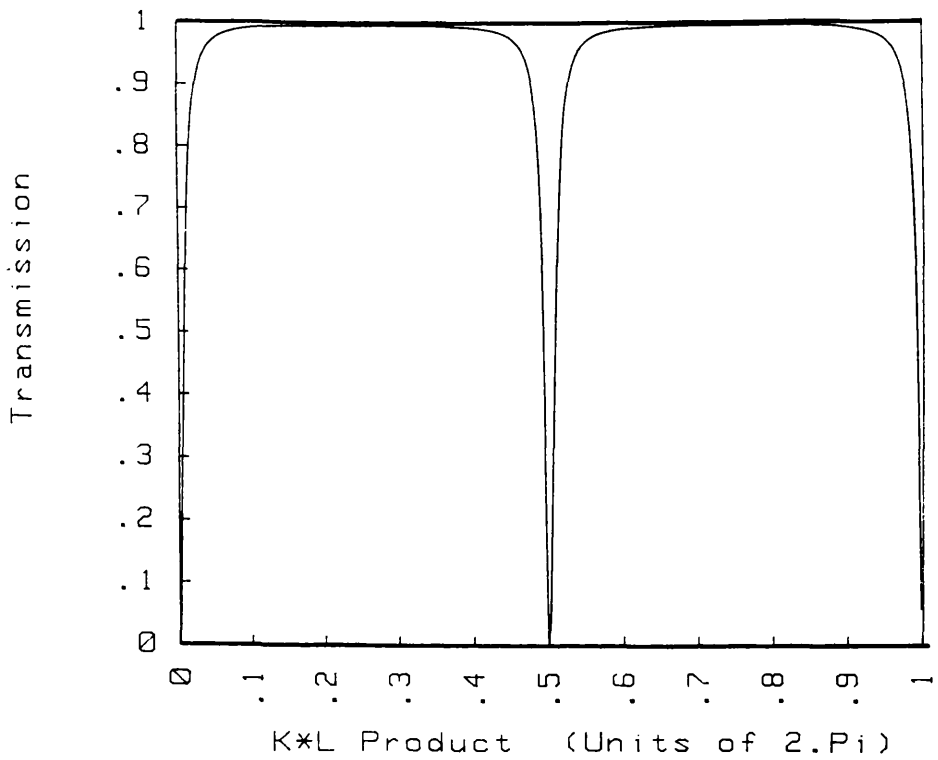


Figure 5.6

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .35 * \Phi_0$

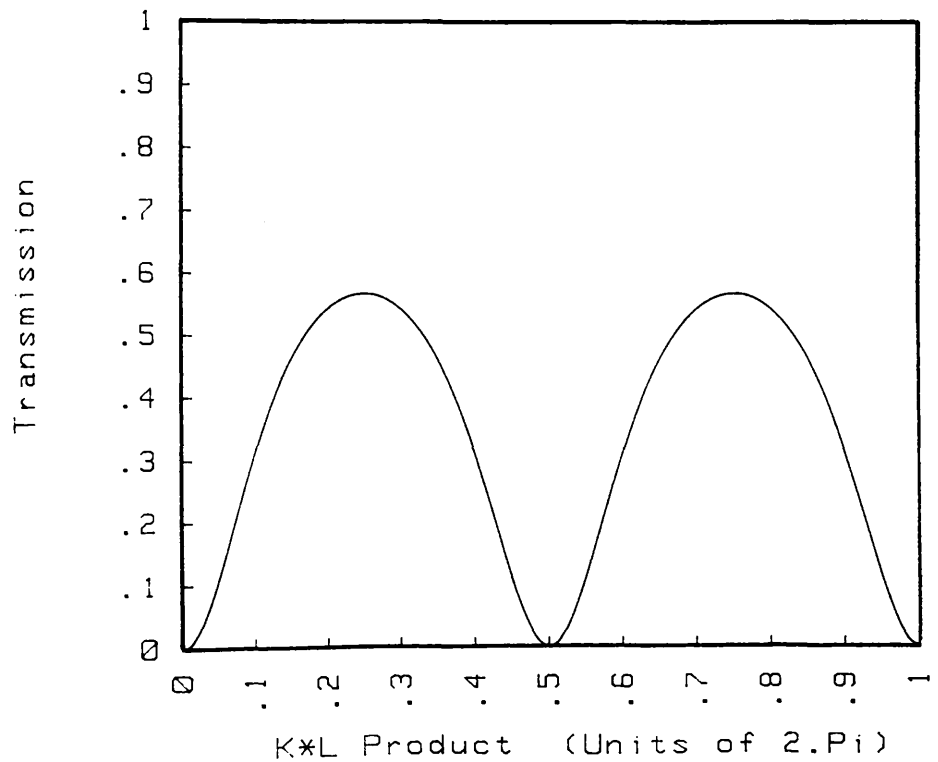


Figure 5.7

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .45 * \Phi_0$

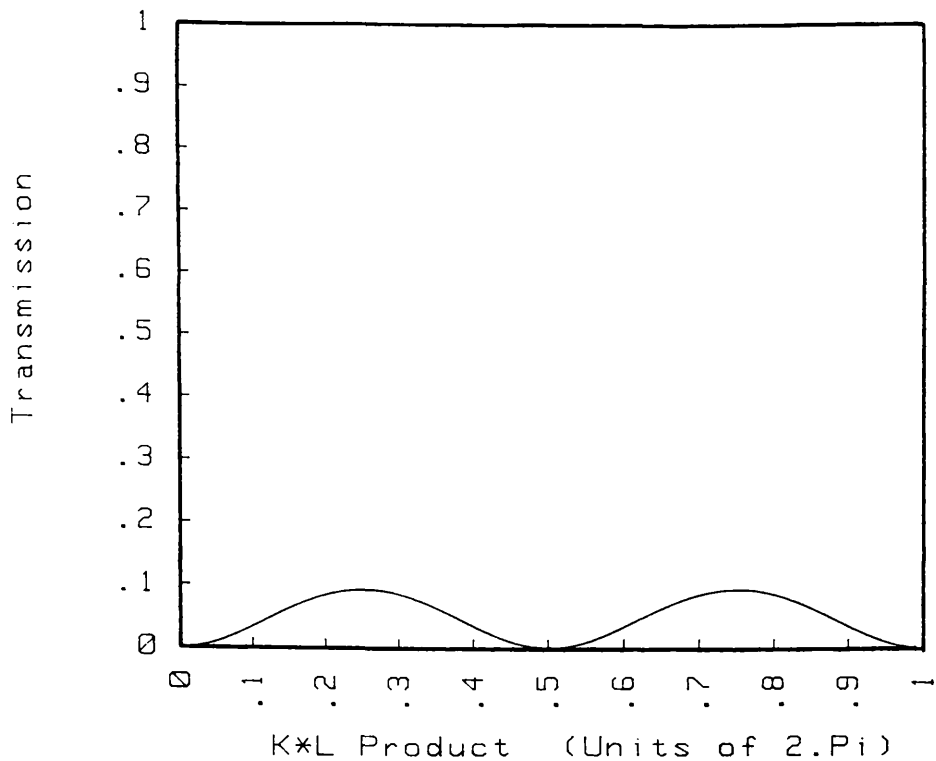


figure 5.8

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .49 * \Phi_0$

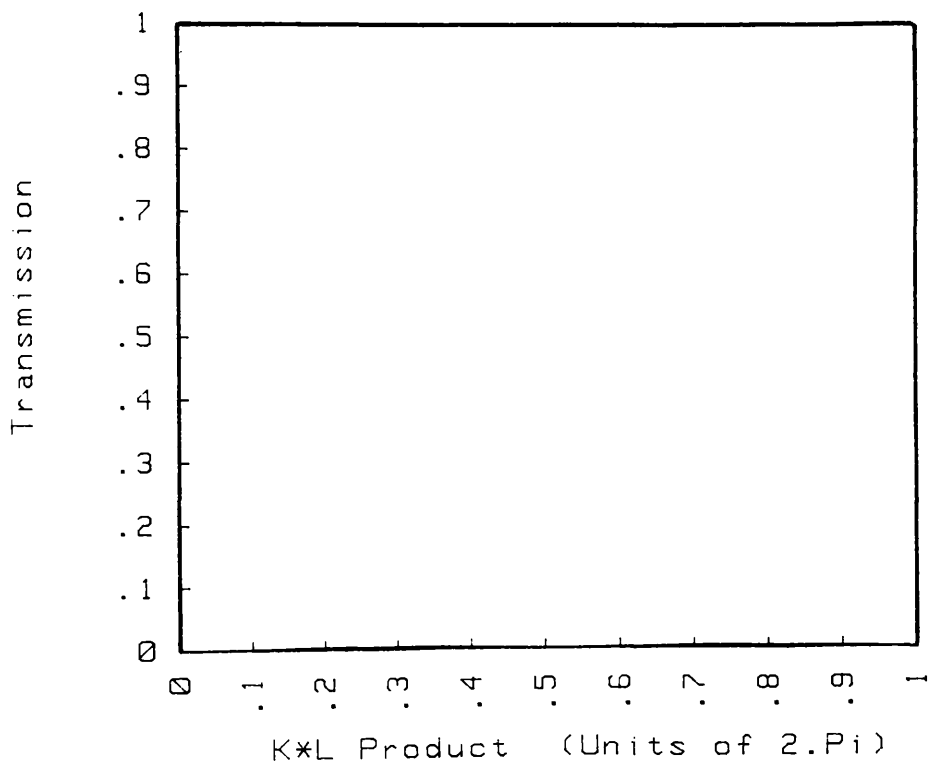


Figure 5.9

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .25 : Flux $\Phi = .1 * \Phi_0$

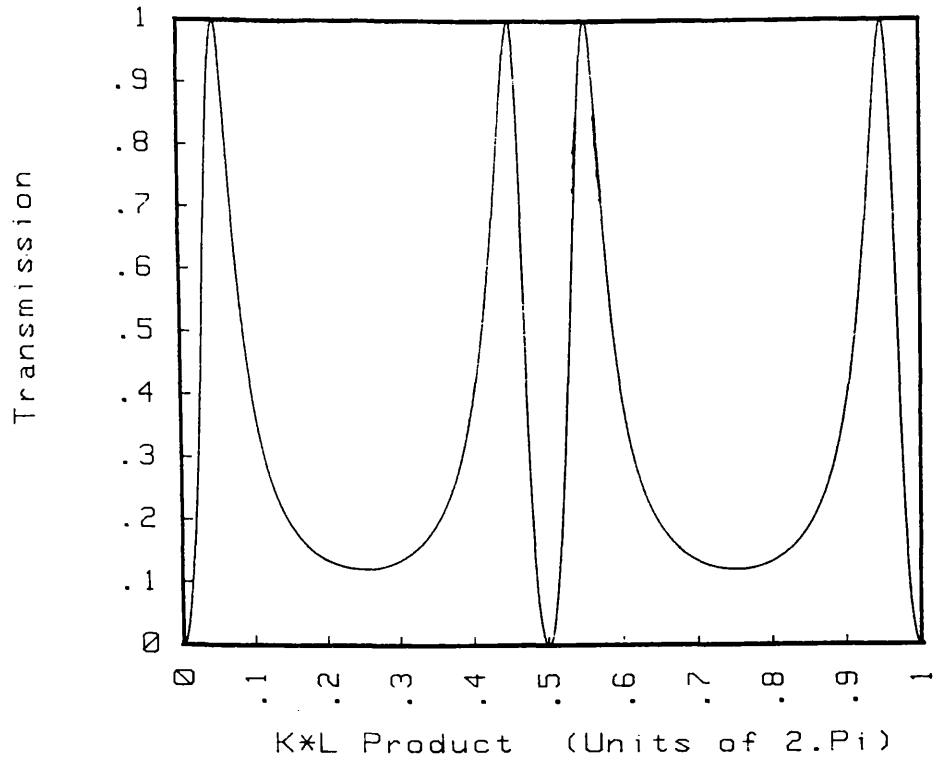


Figure 5.10

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .1 : Flux $\Phi = .1 * \Phi_0$

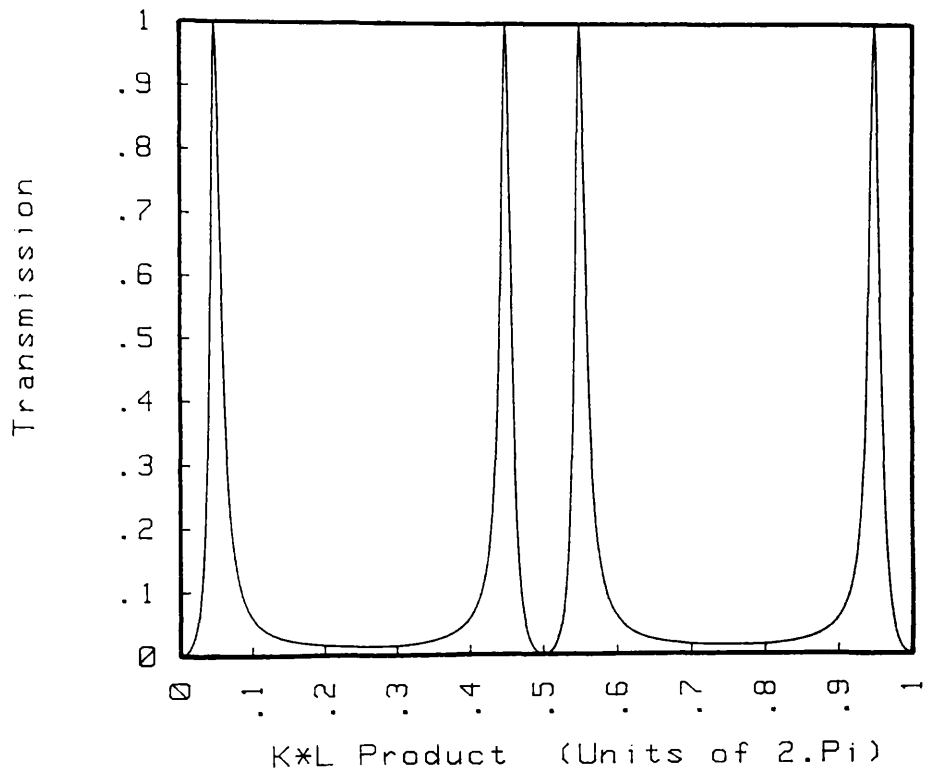


Figure 5.11

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : KL Prod= .1 *2.PI

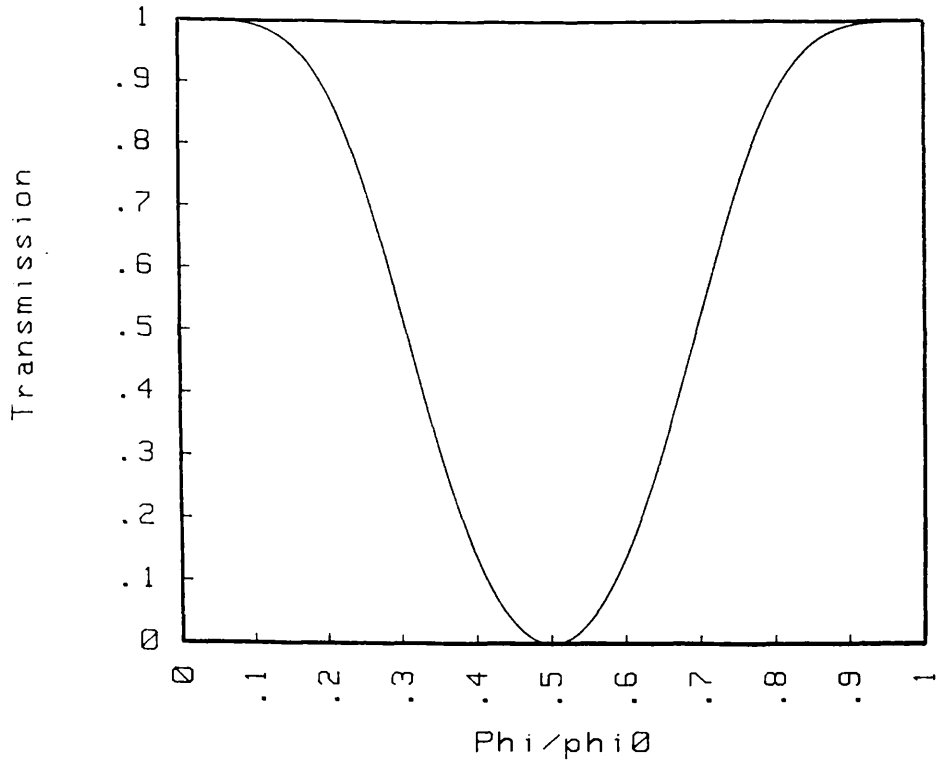


figure 5.12

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : KL Prod= .35 *2.PI

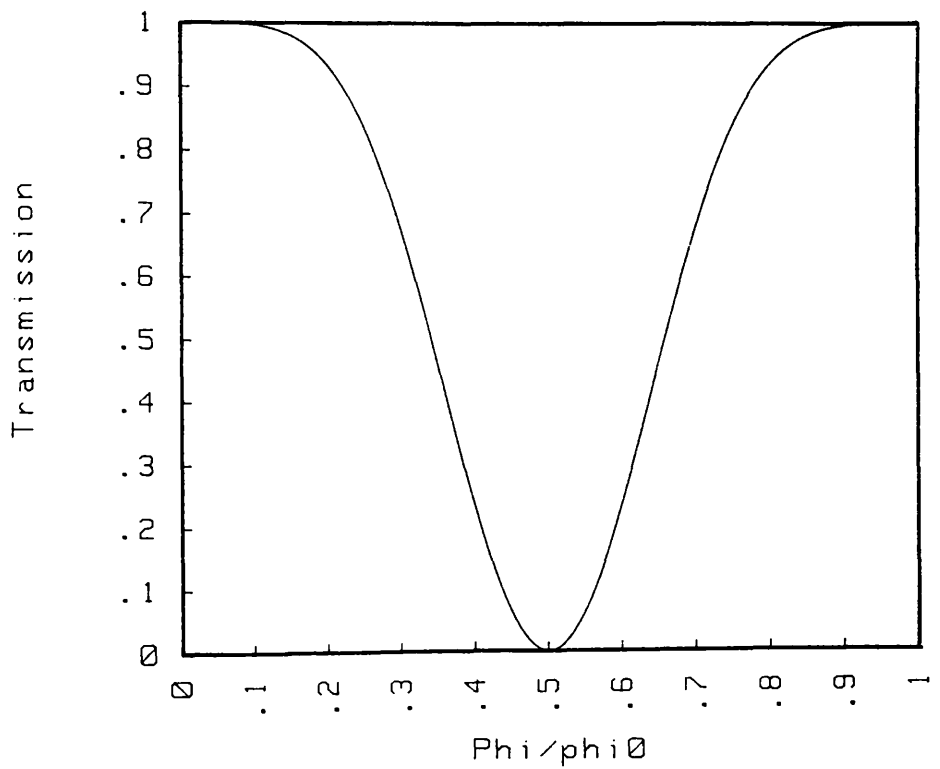


Figure 5.13

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : KL Prod= .45 *2.PI

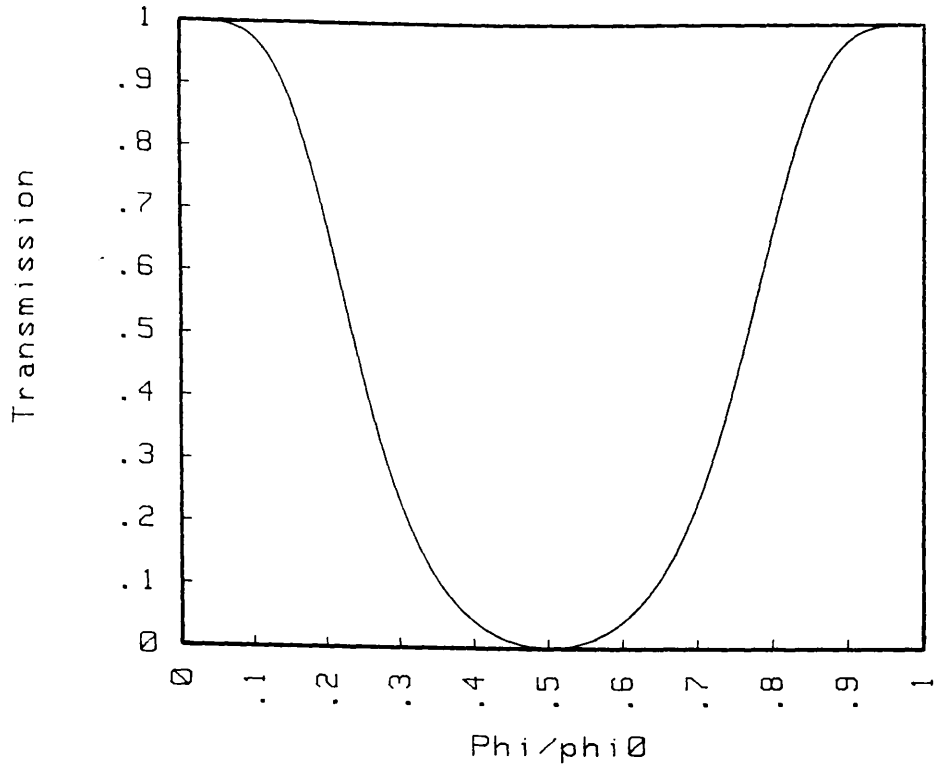


Figure 5.14

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .25 : KL Prod= .35 *2.PI

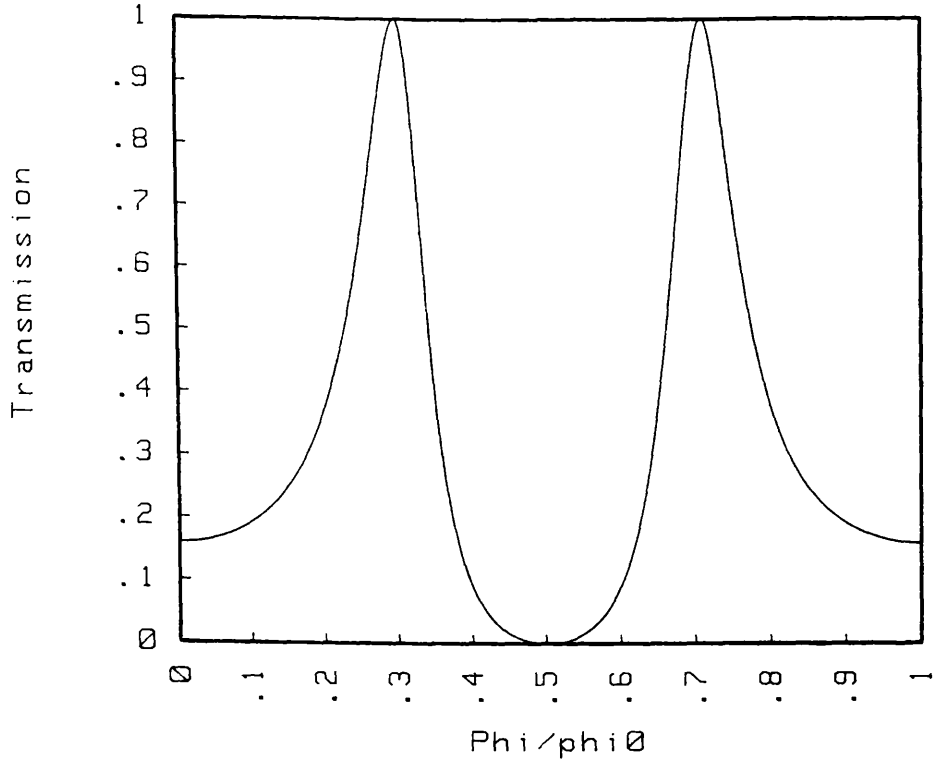


Figure 5.15

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .1 : KL Prod= .35 *2.PI

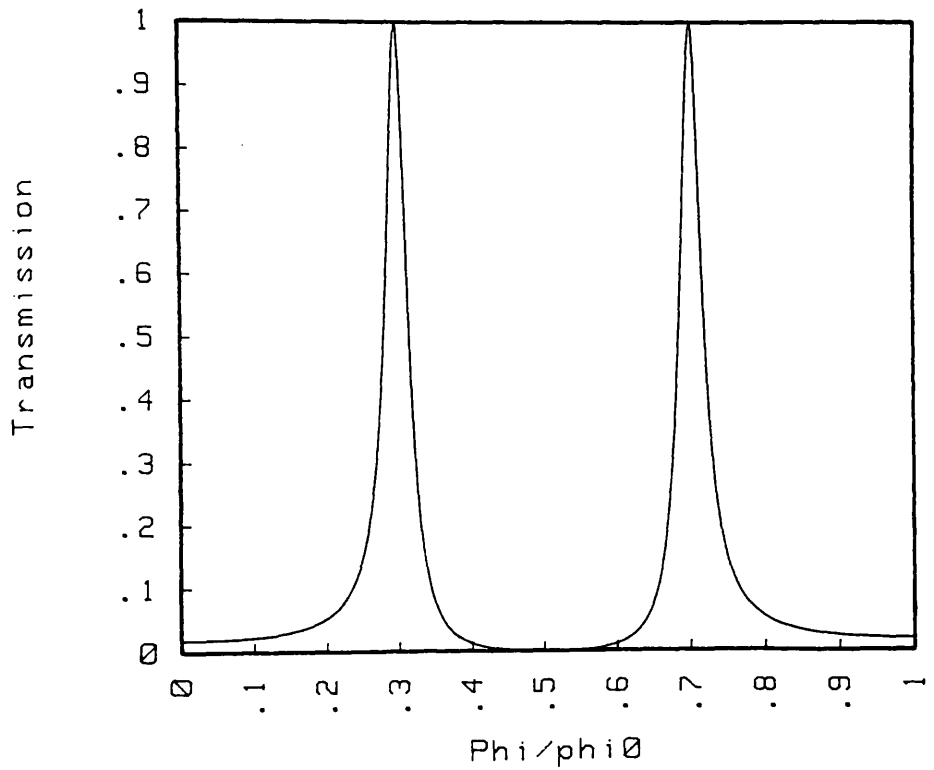


figure 5.16

Figure 5.17

Phasor diagram representation of the wavefunction at the output junction of an AB ring. Zero phase is referenced to the input junction.

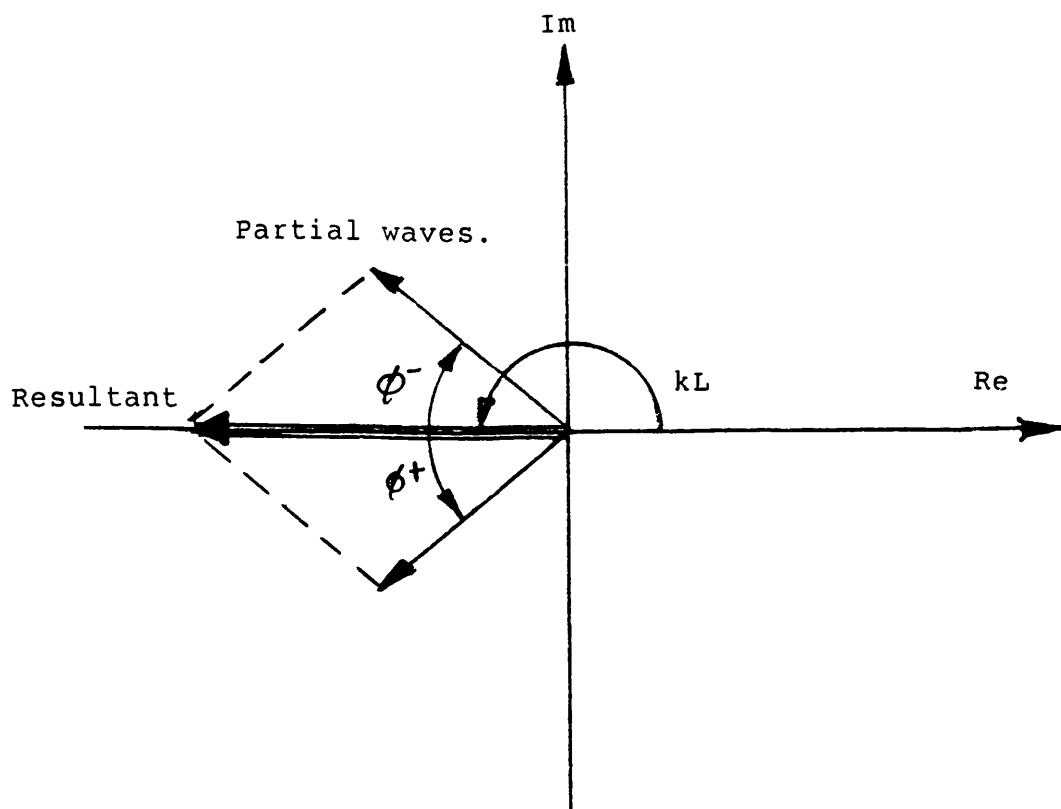


Figure 5.18 Schematic phasor diagram for a ring with a low coupling parameter. The phasor with the thick arrow is the resultant and is generally very small in length.

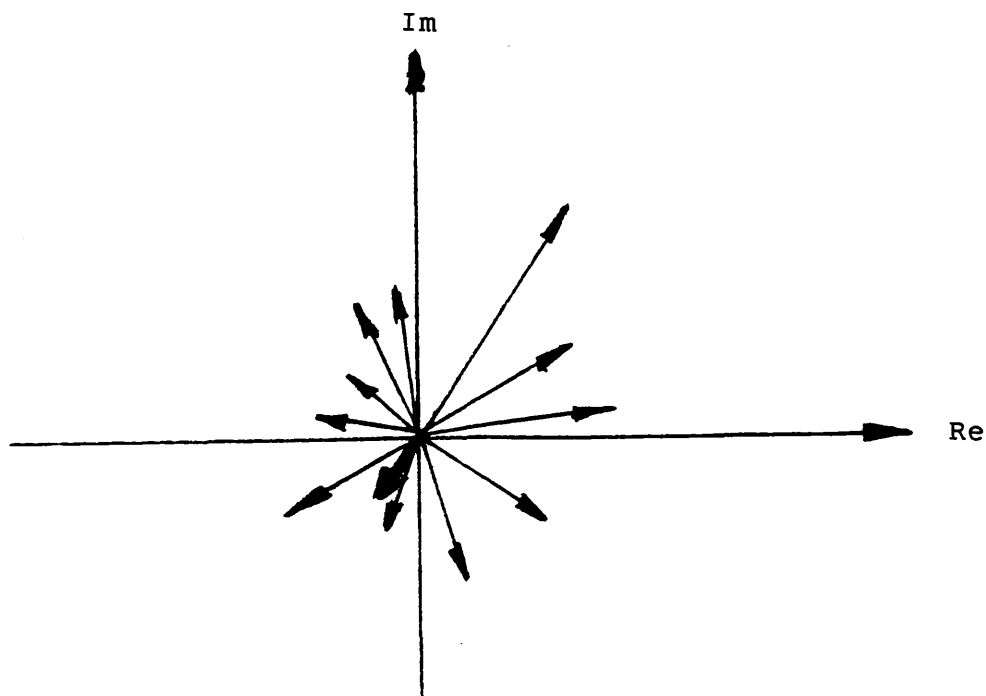


Figure 5.19

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .1 * \Phi_0$

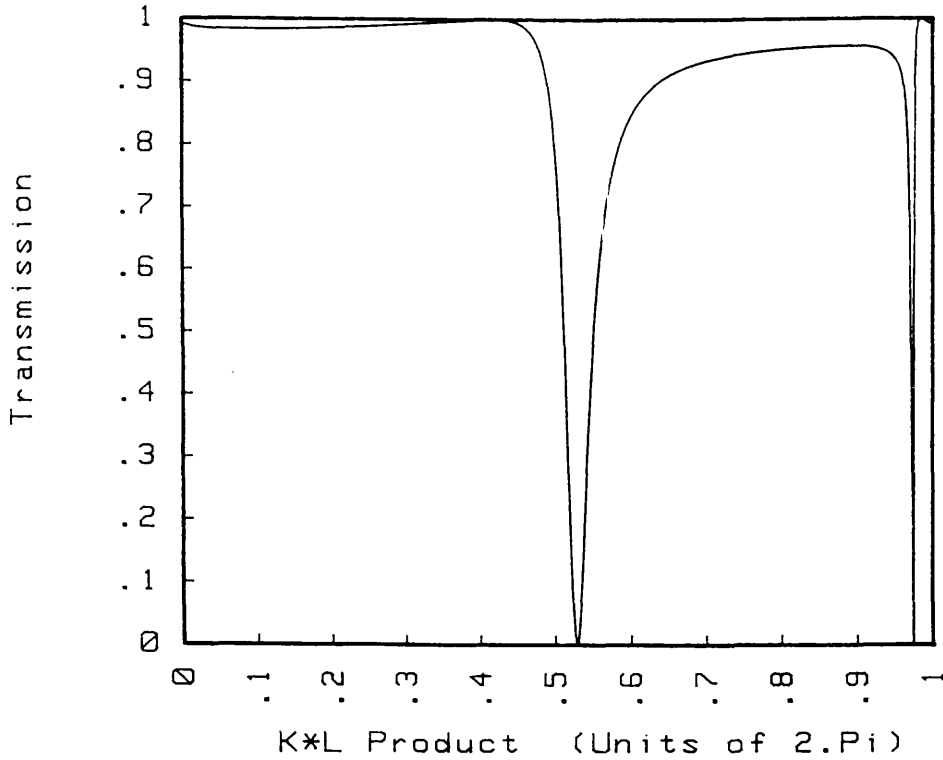


figure 5.20

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .1 * \Phi_0$

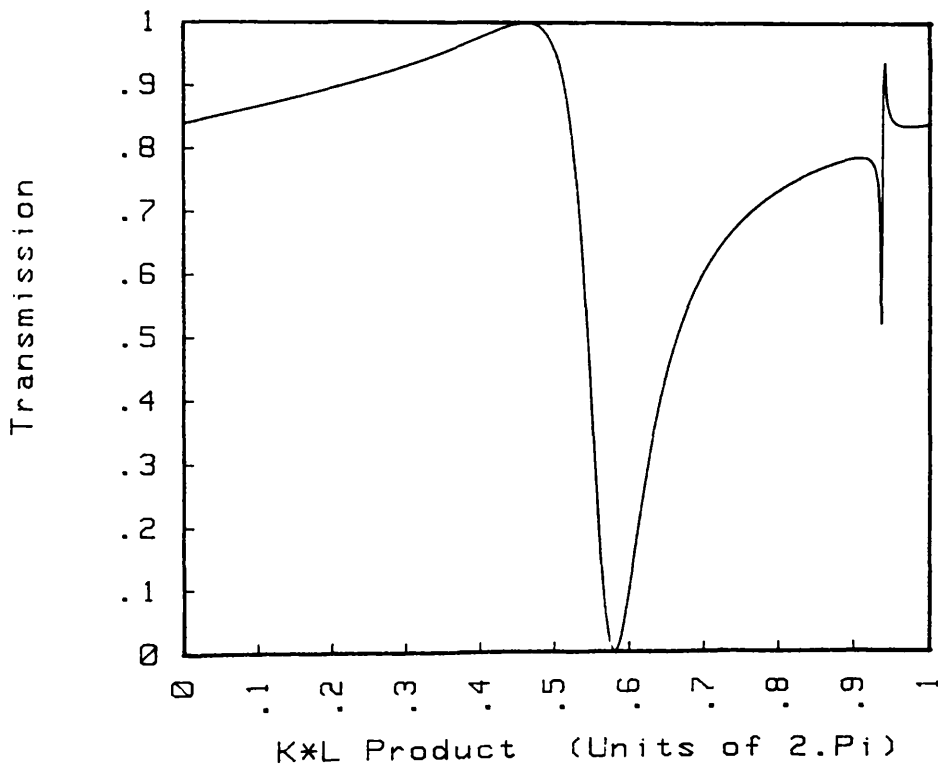


Figure 5.21

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .1 * \Phi_0$

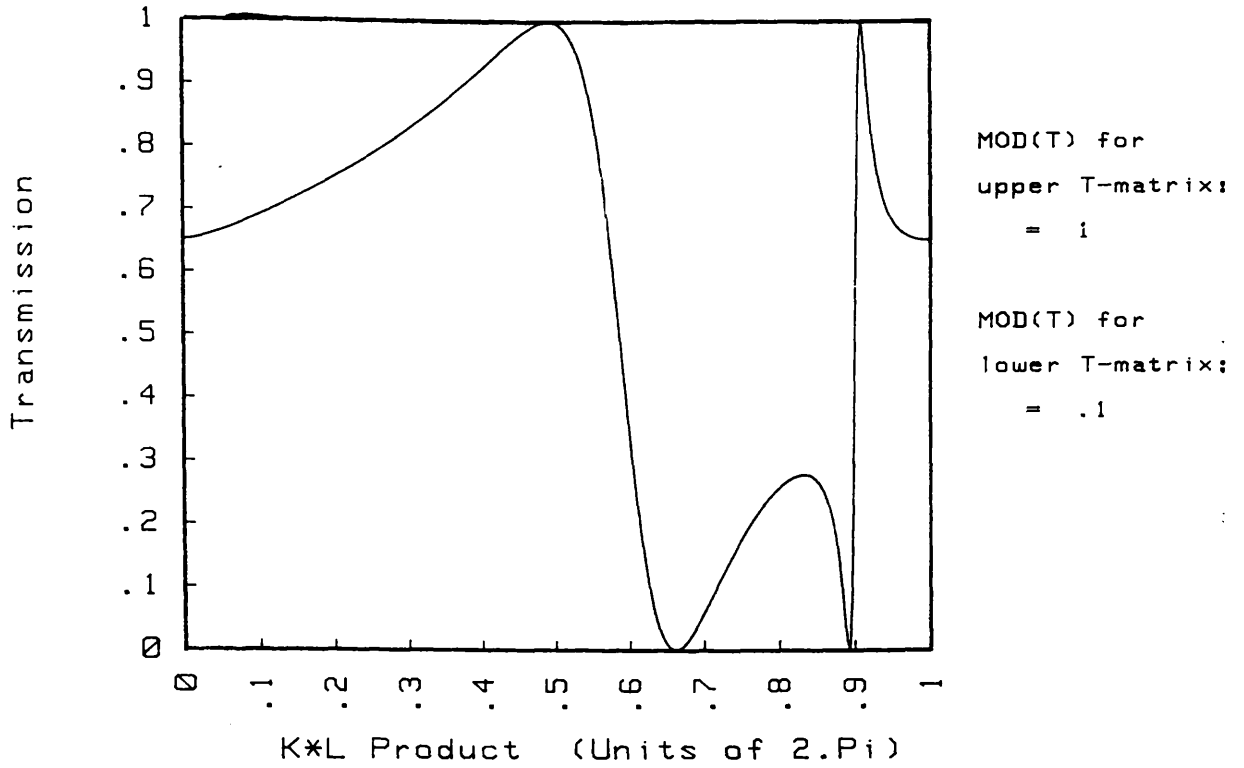


Figure 5.22

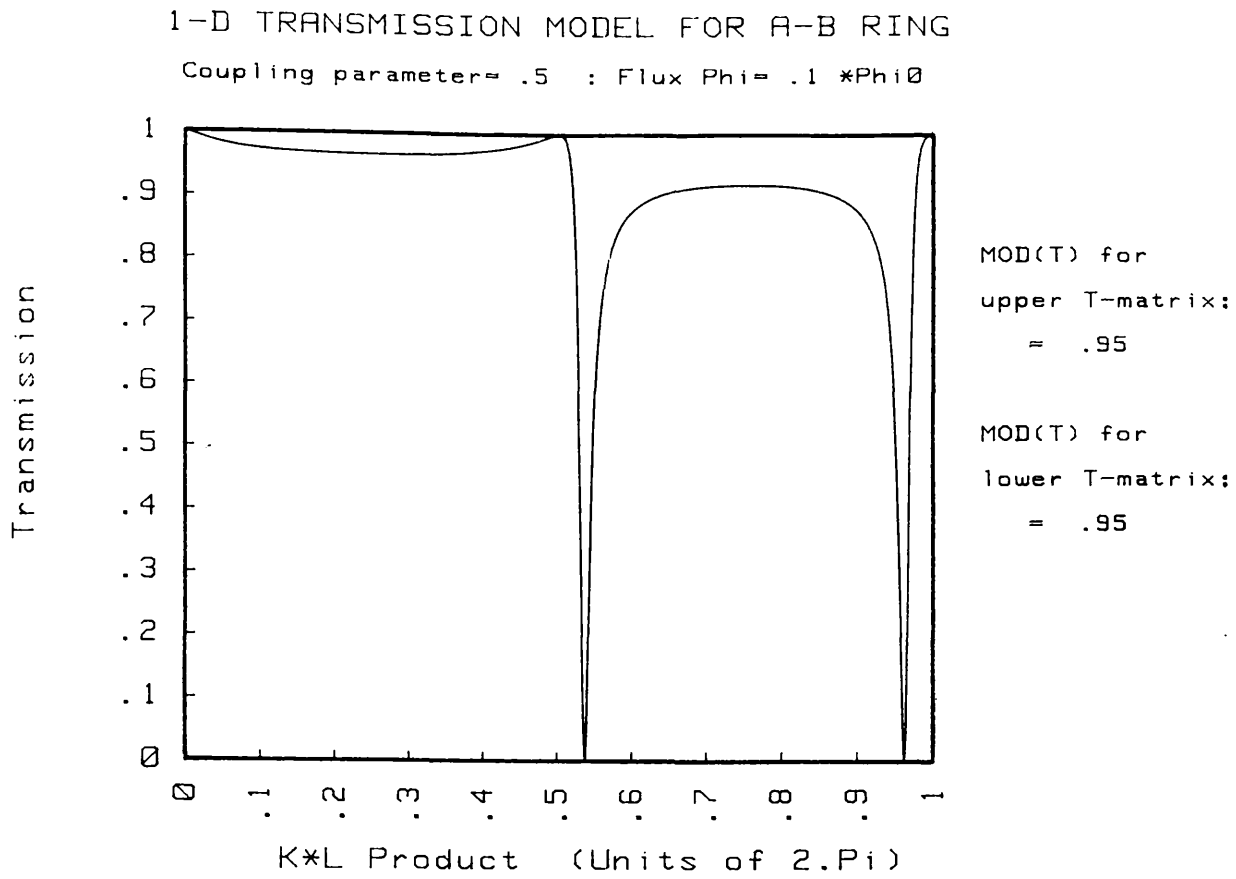


Figure 5.23

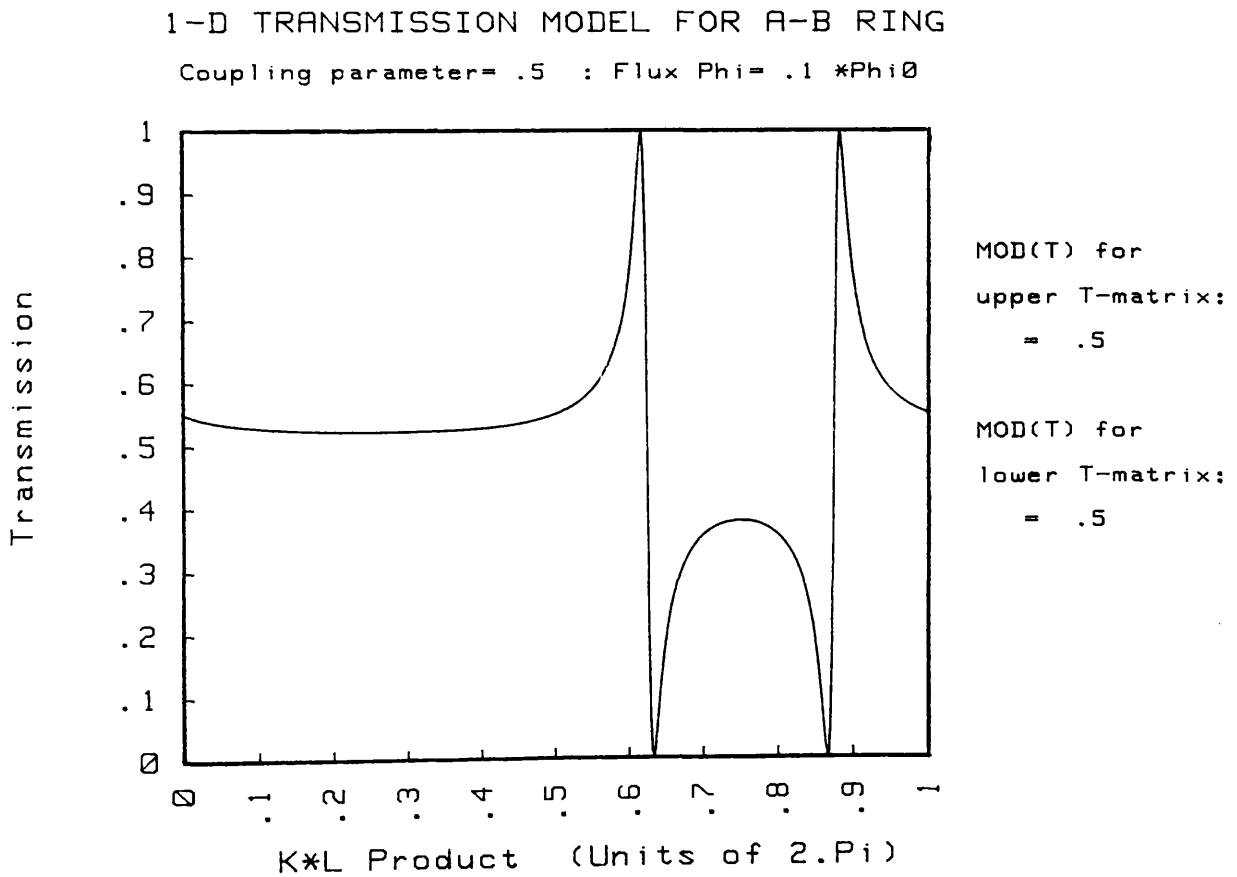
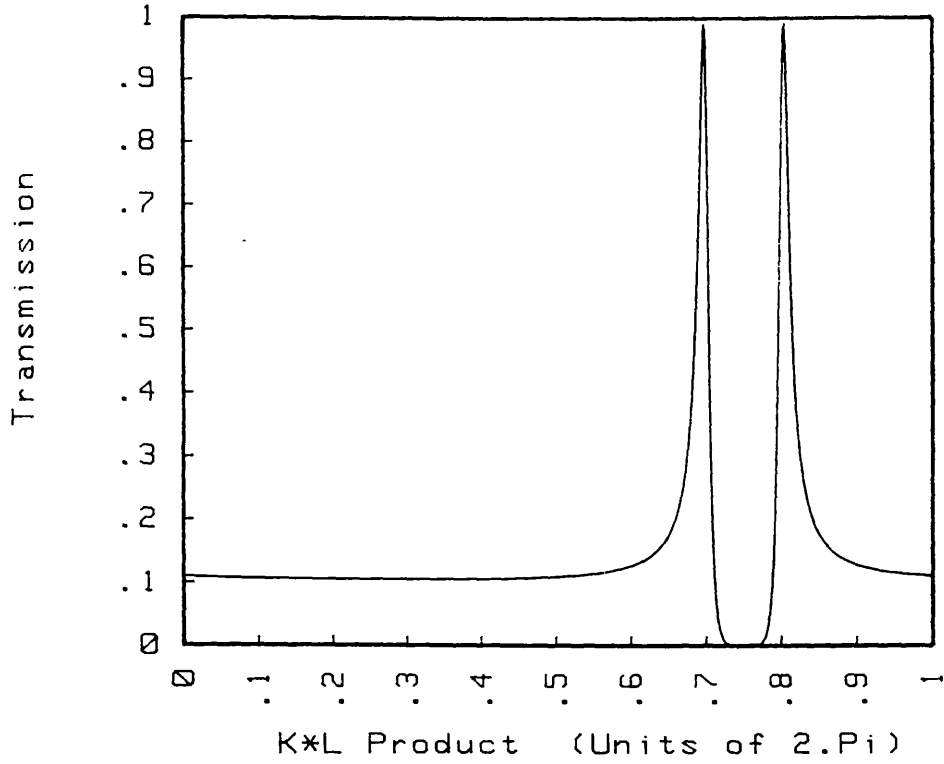


figure 5.24

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .1 * \Phi_0$



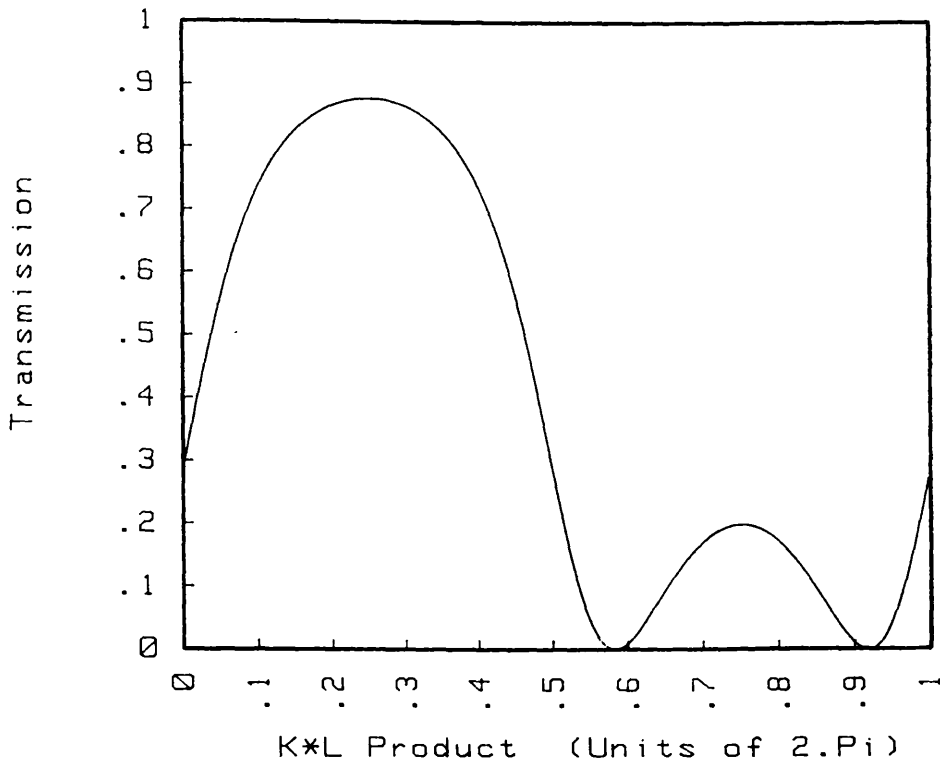
MOD(T) for
upper T-matrix:
= .1

MOD(T) for
lower T-matrix:
= .1

Figure 5.25

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : Flux $\Phi = .35 * \Phi_0$



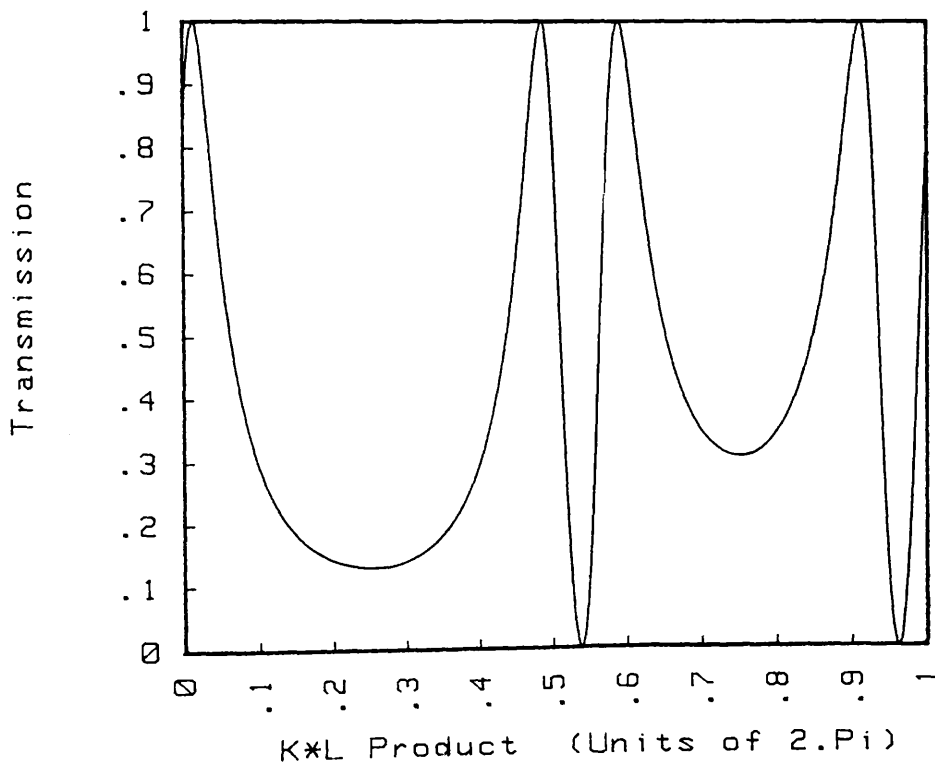
MOD(T) for
upper T-matrix:
= .95

MOD(T) for
lower T-matrix:
= .95

Figure 5.26

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .3 : Flux $\Phi = .1 * \Phi_0$



MOD(T) for
upper T-matrix:
= .95

MOD(T) for
lower T-matrix:
= .95

Figure 5.27

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .1 : Flux $\Phi = .1 * \Phi_0$

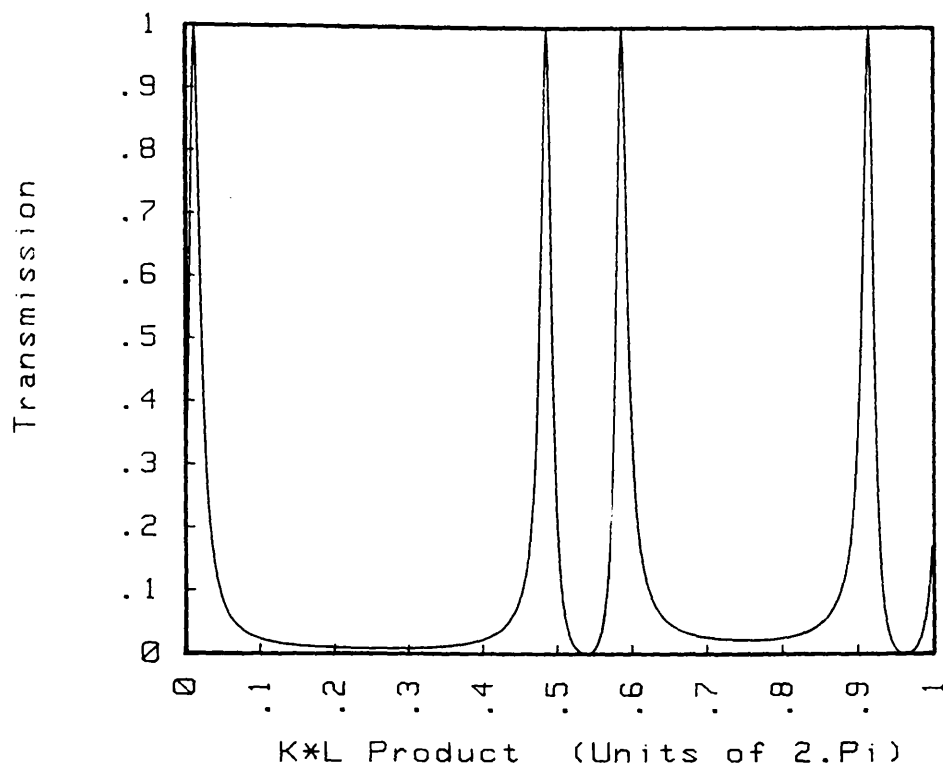
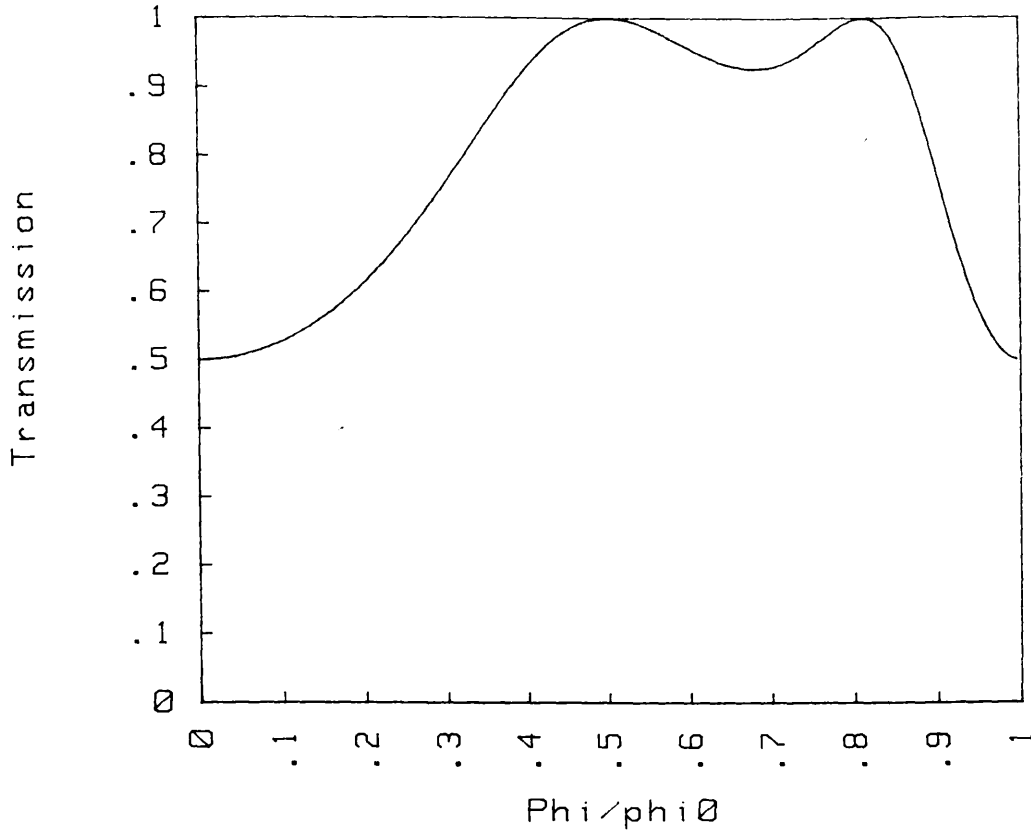


Figure 5.28

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : KL Prod= .1 *2.PI



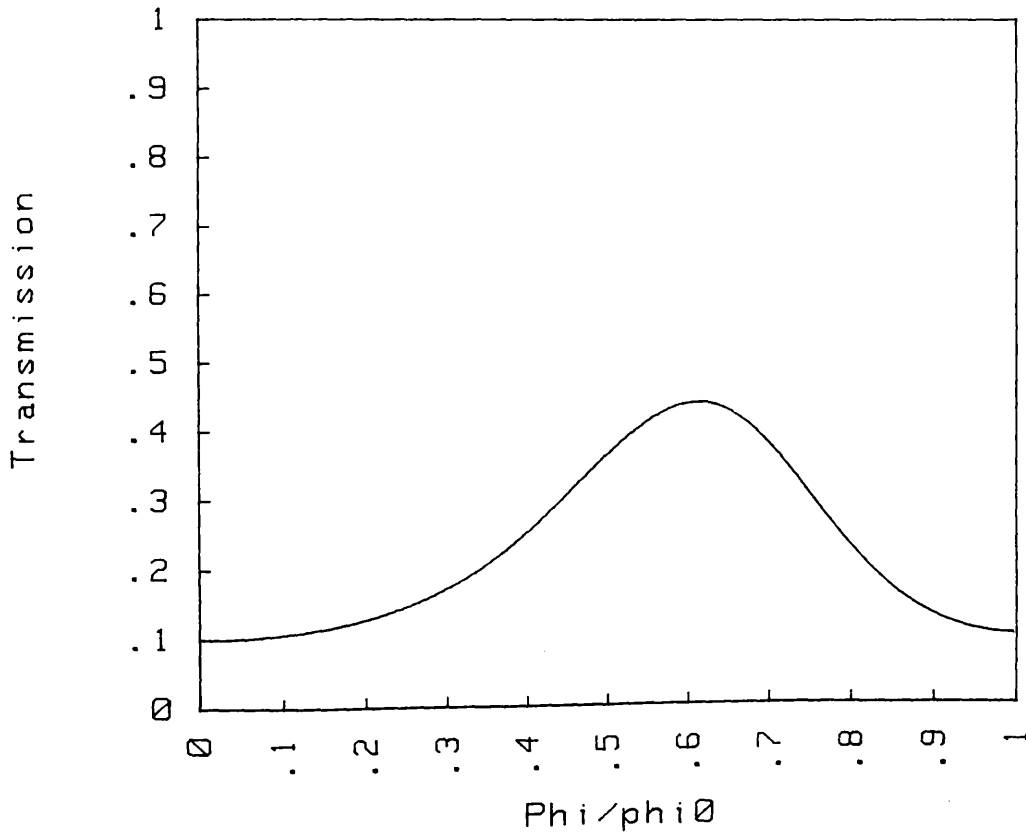
MOD(T) for
upper T-matrix
= .5

MOD(T) for
lower T-matrix
= .5

Figure 5.29

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .5 : KL Prod= .1 *2.PI



MOD(T) for
upper T-matrix
= .1

MOD(T) for
lower T-matrix
= .1

5.7 Summary.

This chapter has described a one-dimensional model which used transmission and scattering matrices to connect plane-wave solutions in a multiply-connected AB ring structure.

The results confirmed the existence of the AB effect in one-dimension and predicted that the transmission should fall to zero when the enclosed flux was equal to half the flux quantum h/e . Also revealed, was the presence of energy resonances, where the transmission may peak at certain values of wavevector due to multiple-reflection processes at the input and output junctions.

Two transmission regimes were identified for the ring. That when the external leads were well coupled to the ring and that when they were poorly coupled. In the absence of scatterers in the arms of ring, the transmission characteristic in the weak-coupling limit was understood in terms of the eigenvalues of a closed ring. In the cases of strong or weak scattering in the weakly-coupled ring and strong scattering in the well-coupled ring, the transmission characteristic was in good agreement with the theory of Buttiker *et al* [1984]. However, the explanation of the structure in the well-coupled ring with no scattering remained rather inconclusive. In this case the electron waves travel ballistically around the ring, a situation which becomes more likely as device dimensions and temperatures are reduced. Since quantum ballistic transport is of central importance to this thesis, it was considered important to develop a theory to understand the behaviour in this regime. To this end, an approach using phasor diagrams was taken as an aid to understanding the transmission plots. It was then shown how the sharp drops in transmission could be explained in the well-coupled regime by considering the resultant phasor at the output junction. This model was also used to show how the sharp peaks could be accounted for in the poorly-coupled regime.

The effect of elastic scattering was investigated by the introduction of transmission matrices whose transmission coefficient was less than unity. The results showed that the symmetry about $kl=0.5*(2\pi)$ for the case with no scattering was destroyed when the arm transmission was lowered. In the case when both arm transmissions were the same, only one symmetry axis remained at $kl=0.75*(2\pi)$. The effects of quite modest scattering ($|T|=0.95$) can have a large effect on the transmission characteristics, in some cases causing a complete reversal of behaviour compared to the no-scattering case.

6 Two-dimensional Time-dependent transport.

6.1 Introduction.

This chapter explains the need for a two-dimensional model of an AB ring and describes the problem in general terms. Discussion is initially confined to a detailed exposition of the solution techniques for the one-dimensional time-dependent problem as this will be later shown to form the basis of the fully two-dimensional computation.

The general numerical problem in one dimension consists of finding a suitable approximant for the time-evolution operator which is both stable and unitary. It is shown that the Cayley form of the time-evolution operator meets both these criteria and that the final numerical problem can be formulated in terms of a tri-diagonal matrix equation for which well-known solution techniques can be used.

The solution strategy for the two-dimensional problem is then introduced. This is based on a decomposition of the total Hamiltonian into two components on which one-dimensional solution techniques can be used, and the problem thus reduced to the solution of many sets of tri-diagonal matrix equations.

One consequence of this procedure is that the general problem cannot be decomposed in this way without introducing a small error in the compound time-evolution operator. Section 6.10 contains an estimate of how serious this is likely to be.

Another problem in solving the two-dimensional problem is that the amount of computer time required is substantially increased over typical 1D problems and therefore attention must be given to the computational efficiency of the algorithms used. Section 6.11 describes how the recursion formulae arising from the solution of the tri-diagonal matrix equations were re-structured to take full advantage of the vector-processing facility on the IBM3090 computer.

6.2 Motivation for a 2D model.

The 1D treatment given in the last chapter had the failing that it predicted a magneto-resistance modulation of 100% of the background value when every experimental value to date was much less than this. A typical value for the modulation depth in metal rings was about 0.1% of the background resistance. Semiconductor structures which utilise the 2DEG at a heterojunction have achieved larger oscillations, typically around 5%-10%, and a maximum of about 20% [Ford 1988]. Despite the improvement in oscillation amplitudes attained over recent years, the 1D model is clearly lacking in its ability to predict the actual oscillation amplitudes.

Another drawback of the 1D model was that it did not take into account the *direct* action of the magnetic field but only the vector potential produced by the field. Thus the phase-shifts in the wavefunction were properly accounted for but no forces acting on the electron were included.

Real rings do not have one-dimensional conductors and the electron wavepacket has some freedom to diffract out of apertures and around structures in two (or three) dimensions. Therefore, in order to model the ring structures in a more realistic fashion, a two-dimensional model was developed. Even though the wavefunction might, in some instances, become appreciably confined in the transverse direction within the wires, it was expected that the inclusion of finite wire widths would help reconcile the difference in MR amplitudes between experiment and the simple 1D model.

6.3 Specification of the problem to be solved.

In recent years, there have been attempts to incorporate quantum effects into conventional Monte Carlo techniques. Of these, the most straightforward is to solve the Schrodinger equation for the transmission and reflection coefficients whenever a particle is incident on a potential barrier (or well). The particle can then be moved across the barrier within a time predicted by the stationary phase approximation. A more elaborate method is based on the evaluation of the Wigner function of the system. In this procedure, a family of Monte Carlo simulations are performed in parallel phase-spaces with different weighting parameters.

It was then shown [Barker and Murray 1983] that the actual process may be reconstructed from a superposition of the family of Monte Carlo simulations. This method however, would consume a (possibly prohibitively) large amount of computer resources. In addition to this problem, a common drawback of Monte-Carlo techniques is that the essential physics can often be obscured by the complexity of the computer model. Hybrid quantum-classical Monte Carlo calculations have been performed [Al-Mudares 1987] for the type of patterned-gate HEMT structures considered here and the results showed that the two-dimensional electron gas within such structures could be successfully confined to form separate narrow conducting channels.

The purpose of this project was to investigate the quantum interference processes occurring in these structures. It was therefore decided to concentrate solely on the quantum transport through a specified structure using a quantum-ballistic transport model, leaving the question of the precise form of the confinement potential as a separate problem.

In this type of model, the motion of the electron is determined by the solution of the time-dependent Schrodinger equation for the system and is considered to move without any collisions (except with the side-walls of the wire); inelastic collisions which change the energy of the electron are not considered. Polar-optic phonon scattering, which is a dominant inelastic collision process at room temperature, ceases to be of importance at the very low temperatures at which experiments are performed. Instead, electron-electron scattering becomes the dominant inelastic process, which proceeds via a number of low-energy transfers [Wind *et al* 1986]. Unfortunately, this process is difficult to incorporate into a simple model, and was neglected in the first instance. Elastic collisions with crystal defects and impurity sites occur more frequently, (although these were also neglected in the first instance) but it was recognised that these processes could be included more easily than the inelastic processes. A qualitative study of impurity scattering is given in section 9.8.

6.4 Solution of the time-dependent Schrodinger equation.

In order to describe the model and solution technique for the 2D case, it is necessary to review the techniques for solving the 1D problem.

It may be shown [Schiff 1955] that a freely-propagating wavepacket has the general form of a plane wave multiplied by a Gaussian envelope. This functional form is known to have a minimal position-momentum uncertainty product. The general 1D problem can therefore be defined as solving the time-dependent Schrodinger equation to find the wavefunction $\psi(x,t)$ in a region where the Gaussian wavepacket is the initial condition $\psi(x,0)$. During the propagation of the wavepacket it may interact with various potentials including single barriers and wells [Goldberg *et al* 1967] and resonant double-barrier structures [Collins 1986].

The Schrodinger equation,

$$H \psi(r,t) = i\hbar \frac{d\psi(r,t)}{dt} \quad (\text{Eqn 6.1})$$

can, in some instances, be efficiently solved by means of an eigenstate expansion of the initial condition. Each eigenstate can be evolved forward in time by multiplying by a phase factor of $\exp(-i\omega t)$, and the total wavefunction reconstructed by a superposition of all such states;

$$\psi(r,t) = \sum a_n \exp(-iE_n t/\hbar) \psi(r) + \int dk a(k) \exp(-iE_n t/\hbar) \psi(r) \quad (\text{Eqn 6.2})$$

where the eigenvectors of the Hamiltonian can, in general, have a part discrete and part continuous spectrum [Messiah 1962]. The eigenvectors satisfy the time-independent equation,

$$H \psi_n(r) = E_n \psi_n(r) \quad (\text{Eqn 6.3})$$

and the expansion coefficients are found by projection of the initial condition onto this basis set:-

$$a_n = \langle \psi_n(r) | \psi(r,0) \rangle \quad (\text{Eqn 6.4})$$

$$a(k) = \langle \psi_k(r) | \psi(r,0) \rangle \quad (\text{Eqn 6.5})$$

This method has two major drawbacks. The first is that for all but the simplest of cases, the eigenvalues and eigenvectors of equation(6.3) must be found numerically, and diagonalising the large Hamiltonian matrices involved can be difficult and computationally intensive. Secondly, when the eigenvalues and eigenvectors have been found, one must then integrate over all eigenvectors with the appropriate expansion coefficients and phase factor (as indicated by equation 6.2) for each spatial point. If there are n^2 points on the grid, then the evaluation of the complete wavefunction at a later time requires $o(n^4)$ operations. For these reasons the above method was not used, and a scheme requiring $o(n^2)$ operations was employed. The difference in computation time between an $o(n^4)$ and an $o(n^2)$ process can be enormous for the grid sizes used. Consider, for example, a 512 x 256 grid and a typical time per floating point operation of 50ns, then an $o(n^2)$ process will take in the order of 6ms compared to about 14 minutes for the $o(n^4)$ process.

The scheme chosen considers how the entire wavefunction can be approximately propagated from one instant in time to another a short time later and relies on the fact that the time-dependent Schrodinger equation can be formally integrated over a time interval Δt to yield:-

$$\psi(r,t+\Delta t) = \exp(-i \Delta t H/\hbar) \psi(r,t) \quad (\text{Eqn 6.6})$$

where the exponential term is known as the time-evolution operator. The Hamiltonian operator appearing in the exponential may be understood in terms of a series expansion.

Naturally, for numerical work, an infinite series expansion of the time evolution operator is not possible. Furthermore the generally complicated nature of the Hamiltonian forces one to seek approximations to the time evolution operator which are as simple as possible. The next few sections consider what criteria must be applied in choosing an expansion scheme and to what extent approximants to the full time-evolution operator fulfil these requirements.

6.5 The stability of expansion schemes for the time evolution operator.

The simplest approximation to the time-evolution operator is a first order Taylor expansion, which results in the following explicit scheme,

$$\psi(r)^{n+1} = (1 - iH\Delta t/\hbar) \psi(r)^n \quad (\text{Eqn 6.7})$$

The scheme is termed "explicit" because the wavefunction at timestep $n+1$ is found directly from the wavefunction at timestep n . The stability of this scheme is investigated by the use of the Von Neumann stability criterion [Potter 1973] which considers possible unbounded growth in any of the Fourier modes present on the mesh of the form:-

$$u_j^n = u^n \exp(ikx_j) \quad (\text{Eqn 6.8})$$

for the k^{th} Fourier mode, where n denotes the timestep index and j the meshpoint index. Direct substitution into the discretised form of the explicit scheme shows how each of the modes will propagate.

$$u^{n+1} = (1 - i(2A(1 - \cos(k\Delta x)) + V_j\Delta t)) u^n, \quad A = \Delta t / (2\Delta x^2) \quad (\text{Eqns 6.9})$$

The growth factor, defined as $G = u^{n+1}/u^n$, for a Fourier mode of wavevector k is therefore:-

$$1 - i(2A(1 - \cos(k\Delta x)) + V_j\Delta t) \quad (\text{Eqn 6.10})$$

The Von Neumann stability requirement states that for a discretisation scheme to be stable the modulus of the growth factor (to first order terms in Δt) must be less than or equal to unity. The above growth factor is greater than unity for all wavenumbers, and so the explicit scheme is inherently unstable. A simple explicit method is therefore not suitable for the solution of the time-dependent Schrodinger equation.

The next approximation considered was the Crank-Nicholson scheme. This scheme considers forward and backward propagation in time from the current wavefunction:-

$$\psi^{n+1} = \exp(-i\Delta t H/\hbar) \psi^n \quad (\text{Eqn 6.11})$$

$$\psi^{n-1} = \exp(+i\Delta t H/\hbar) \psi^n \quad (\text{Eqn 6.12})$$

and uses the two relations to eliminate ψ^n , resulting in the equation:-

$$\exp(+i\Delta t H/\hbar) \psi^{n+1} = \exp(-i\Delta t H/\hbar) \psi^{n-1} \quad (\text{Eqn 6.13})$$

Taylor-expansion of the two exponentials to first order in time results in the approximation:-

$$\psi^{n+1} = \left((1 - i(\Delta t/2) H/\hbar) / (1 + i(\Delta t/2) H/\hbar) \right) \psi^n \quad (\text{Eqn 6.14})$$

This approximate form of the time-evolution operator is known as the Cayley expansion.

A stability analysis similar to that performed for the explicit scheme yields a growth factor of:-

$$G = (1 - i(2A(1 - \cos(k\Delta x)) + 2V_j\Delta t)) / (1 + i(2A(1 - \cos(k\Delta x)) + 2V_j\Delta t)) \quad (\text{Eqn 6.15})$$

the modulus of which is always equal to unity, which means that errors will not grow exponentially. It also has the desirable property that it is accurate to second order in both Δx and Δt [Potter 1973].

6.6 The importance of unitarity in the expansion scheme.

In choosing a numerical integration scheme, one must not only consider the growth of any errors, but also the unitarity of the scheme. In physical terms the property of unitarity applied to the time-evolution operator ensures that the overall normalisation of the wavefunction remains a constant and hence the number of "particles" in the simulation is conserved.

This can be seen by considering the wavefunction as a vector in Hilbert space, with the co-ordinate axes the eigenvectors of the Hamiltonian matrix. Since the eigenvectors form a complete set and are orthogonal to each other, the norm of the wavefunction can be written as:-

$$\text{Norm} = \sum_n a_n \langle \psi_n | \psi_n^* \rangle \quad (\text{Eqn 6.16})$$

where the a_n are the eigenvector expansion coefficients. This shows that the normalisation may be thought of as the length of the state vector. A unitary matrix has the general property of being able to rotate the state vector (*i.e.* propagate the solution in time) without altering its length and so will maintain the normalisation of the wavefunction.

An operator is unitary if it satisfies the identity:-

$$\mathbf{O}^\dagger \mathbf{O} = 1 \quad (\text{Eqn 6.17})$$

where \mathbf{O}^\dagger is the Hermitian conjugate of the operator \mathbf{O} . Making use of the fact that the Hamiltonian operator is an Hermitian operator, it is seen that the exact form of the time-evolution operator is exactly unitary, as expected.

$$\begin{aligned} \mathbf{L}^\dagger \mathbf{L} &= \exp(+i\mathbf{H}^\dagger \Delta t/\hbar) \exp(-i\mathbf{H} \Delta t/\hbar) \\ &= 1, \quad (\mathbf{H}^\dagger = \mathbf{H}) \end{aligned} \quad (\text{Eqn 6.18})$$

Applying the same test to the explicit expansion scheme results in:-

$$\begin{aligned} \mathbf{L}^\dagger \mathbf{L} &= (1+i\Delta t \mathbf{H}^\dagger/\hbar) (1-i\Delta t \mathbf{H}/\hbar) \\ &= 1 + \Delta t^2 \mathbf{H}^2/\hbar^2 \end{aligned} \quad (\text{Eqn 6.19})$$

so the explicit scheme, in addition to being unstable is also non-unitary. For the Crank-Nicholson scheme, the unitarity test results in:-

$$\begin{aligned} \mathbf{L} \mathbf{L}^\dagger &= ((1-i\Delta t \mathbf{H}/\hbar) / (1+i\Delta t \mathbf{H}/\hbar)) ((1+i\Delta t \mathbf{H}/\hbar) / (1-i\Delta t \mathbf{H}/\hbar)) \\ &= 1 \end{aligned} \quad (\text{Eqn 6.20})$$

since $(1-i\Delta t \mathbf{H}/\hbar)$ commutes with $(1+i\Delta t \mathbf{H}/\hbar)$. Thus, the Crank-Nicholson scheme is exactly unitary and so conserves the normalisation of the wavefunction.

6.7 Numerical solution of the equations.

Because of the excellent norm-preserving properties and good stability, the Crank-Nicholson scheme was chosen as the integration method in this project. However, one drawback in using the Crank-Nicholson scheme is that it results in a set of implicit equations which means that the wavefunction at a particular point cannot be found directly from the previous timestep because it is coupled to the solution at surrounding points for the next timestep. The wavefunction at each point must therefore be determined by the solution of a system of equations. To illustrate this point, consider a simple one-dimensional Hamiltonian of the form:-

$$H = (-\hbar^2/2m^*) d^2/dx^2 + V(x) \quad (\text{Eqn 6.21})$$

The discretised Hamiltonian acting on the wavefunction at a particular timestep then becomes:-

$$(-\hbar^2/2m^*) (\psi^n(p+1) - 2\psi^n(p) + \psi^n(p-1)) / \Delta x^2 + V(p)\psi^n(p) \quad (\text{Eqn 6.22})$$

where p is the mesh-point index and n is the timestep index. The centred form of differencing scheme has been used here because Taylor expansion shows that this is correct to order Δx^2 instead of only Δx for the one-sided schemes. There is also the bonus of a neater, more symmetrical-looking set of equations when the Hamiltonian is later split for two-dimensional computations.

In one-dimension, the Crank-Nicholson scheme expands to read:-

$$(1 - i\alpha (\psi^{n+1}(p+1) + \psi^{n+1}(p-1) - U(p) \psi^{n+1}(p))) = (1 + i\alpha (\psi^n(p+1) + \psi^n(p-1) - U(p) \psi^n(p))) \quad (\text{Eqn 6.23})$$

where,

$$\alpha = \hbar \Delta t / (4m^* \Delta x^2), \quad U(p) = \Delta x^2 / 2m^* V(p) / \hbar^2 + 2 \quad (\text{Eqns 6.24})$$

Equation 6.23 may then be simplified and recast as a tri-diagonal matrix equation:-

$$a(p)\psi^{n+1}(p+1) + b(p)\psi^{n+1}(p) + c(p)\psi^{n+1}(p-1) = d(p) \quad (\text{Eqn 6.25})$$

where $a(p)$, $b(p)$ and $c(p)$ are complex constants which can be explicitly computed, but in general change from one mesh point to another. The quantity $d(p)$ is to be called the source term since it has the same mathematical form as a source charge would in a discretised solution of Poisson's equation. The source term may be computed explicitly from the current wavefunction using the right hand side of equation 6.23.

6.8 Solution of the tri-diagonal matrix equations.

A standard Gaussian elimination of N unknowns in an $N \times N$ matrix requires $O(N^3)$ arithmetic operations. The tri-diagonal system however, has a particularly simple solution which requires only $O(N)$ operations. Following standard theory, it is assumed that $\psi(p)$ satisfies the following recurrence relation:-

$$\psi(p) = g(p) - \omega(p) \psi(p+1) \quad (\text{Eqn 6.26})$$

Substituting this relation into the tri-diagonal matrix equation and re-arranging yields the recurrence formula.

$$a(p)\psi(p-1) = (d(p) - c(p)g(p)/\omega(p)) - (b(p) - c(p)/\omega(p)) \psi(p) \quad (\text{Eqn 6.27})$$

Since $a(p)$, $b(p)$, $c(p)$ and $d(p)$ are known at each point, the value of the wavefunction at point $n-1$ can be found from the wavefunction at point n if the auxiliary factors $g(p)$ and $\omega(p)$ can be determined. The auxiliary factors can also be found recursively by comparing the above equation with the original assumed form in equation 6.26. The recurrence relations are:-

$$g(p-1) = (d(p) - c(p)g(p)/\omega(p))/a(p) \quad (\text{Eqn 6.28})$$

$$\omega(p-1) = (b(p) - c(p)/\omega(p))/a(p) \quad (\text{Eqn 6.29})$$

which can be re-arranged to give the forward recursion formulae:-

$$\omega(p) = c(p)/(b(p) - a(p) \omega(p-1)) \quad (\text{Eqn 6.30})$$

$$g(p) = \omega(p) (d(p) - a(p) g(p-1))/c(p) \quad (\text{Eqn 6.31})$$

The general strategy for solving the equations therefore, is to first calculate the auxiliary variables $\omega(p)$ and $g(p)$ in a forward sweep of the mesh from points 1 to N followed by a backward sweep to find the solution using equation 6.26 and the previously determined values of $\omega(p)$ and $g(p)$. For computer implementation it is useful to note that $\omega(p)$ can be substituted into the expression for $g(p)$, resulting in equal denominators for the two recurrence relations. The denominator may therefore be calculated only once for both functions, saving the number of arithmetic operations required.

The boundary conditions are imposed by specifying the values of $\psi(0)$ and $\psi(N+1)$ at the edges of the mesh. The value of $\psi(0)$ is used for initiating the forward recurrence:-

$$b(0)\psi(0) = d(0) - c(0) \psi(1) \quad (\text{Eqn 6.32})$$

so that

$$g(1) = (d(1) - a(1) \psi(0))/b(1) \quad (\text{Eqn 6.33})$$

and

$$\omega(1) = c(1)/b(1) \quad (\text{Eqn 6.34})$$

The value of $\psi(N+1)$ defines the starting value for the back-substitution stage.

$$\psi(N) = g(N) - \omega(N) \psi(N+1) \quad (\text{Eqn 6.35})$$

In this work, the behaviour of the wavefunction will be considered to be occurring in a closed system and the boundary conditions used will be :-

$$\psi(0)=0, \quad \psi(N+1) = 0 \quad (\text{Eqns 6.36})$$

6.9 Generalisation to two dimensions.

As explained in section 6.4 the method of expansion in eigenfunctions is not well suited for numerical computation due to the very large number of arithmetic operations required for any reasonable size mesh.

However in two dimensions, the direct use of the full two-dimensional Hamiltonian in the Cayley expansion scheme also leads to a large computational problem. The resulting matrix equation to be solved has a banded structure but it is not tri-diagonal and so the very-efficient tri-diagonal algorithm described in section 6.8 cannot be used. Through the use of sparse matrix techniques the number of operations required can be expected to be less than $O(N^3)$ needed for a full Gaussian elimination but will nevertheless rise sharply with increasing mesh size.

Instead of dealing directly with the two-dimensional Hamiltonian therefore, a strategy for solving the problem by reduction to many sets of 1D problems was sought. If this could be achieved the efficient techniques already discussed could be used, resulting in an $O(N)$ solution time.

The starting point is to resolve the Hamiltonian into two components labelled x and y :-

$$H(\text{tot}) = H_x + H_y \quad (\text{Eqn 6.37})$$

although it must be stressed that this notation is not meant to indicate that the two portions are solely functions of x or y .

Substituting this into the Cayley expansion yields the relation:-

$$\psi^{n+1} = (1 - H_x \tau - H_y \tau) (1 + H_x \tau + H_y \tau)^{-1} \psi^n, \quad \tau = i\Delta t / 2\hbar \quad (\text{Eqn 6.38})$$

This expression may then be factored so that it resembles the product of two one-dimensional time-evolution operators acting on the current wavefunction:-

$$\psi^{n+1} = ((1 - H_x \tau) (1 - H_y \tau)) / ((1 + H_x \tau) (1 + H_y \tau)) \psi^n \quad (\text{Eqn 6.39})$$

which, by comparison with equation 6.38, is correct to terms of order Δt^2 .

If an intermediate wavefunction $\psi^{n+1/2}$ is introduced as an auxiliary variable, the calculation can be broken-down into two stages:-

$$\psi^{n+1/2} = ((1 - H_y \tau) / (1 + H_x \tau)) \psi^n \quad (\text{Eqn 6.40})$$

$$\psi^{n+1} = ((1 - H_x \tau) / (1 + H_y \tau)) \psi^{n+1/2} \quad (\text{Eqn 6.41})$$

to arrive at two equations of one-dimensional form. This procedure is known as the alternating direction implicit (ADI) method and is frequently used for the solution of the closely-related time-dependent heat flow equation in two dimensions.

This then is the solution strategy chosen for the two-dimensional time-dependent Schrodinger equation. The following chapters will give explicit forms for the Hamiltonian operator and the details of how the formulae were implemented in a computer model will be discussed.

6.10 Errors arising from the non-commutivity of the split-operators.

In the last section it was shown how the Cayley expansion for the time-evolution operator using the full Hamiltonian could be formulated as two separate equations of one-dimensional form. However, in introducing the half-timestep wavefunction and performing the calculation in two distinct stages another error in addition to those caused by the finite expansion of the time-evolution operator has been introduced which arises from the non-commutivity of the L_x and L_y operators. The Schrodinger equation can be written in simplified form as:-

$$d\psi/dt = -L \psi \quad (\text{Eqn 6.42})$$

And this equation can then be integrated over a timestep interval Δt to give the solution at timestep $n+1$ from the wavefunction at timestep n .

$$\psi^{n+1} = \exp(-L\Delta t) \psi^n \quad (\text{Eqn 6.43})$$

The ADI procedure, described in section 6.9, then required splitting the total operator into two parts. To investigate this procedure, the ADI method is analysed in terms of the unexpanded exponentials.

Resolving the total L operator into two parts as before gives:-

$$\psi^{n+1} = \exp(-(L_x + L_y) \Delta t) \psi^n \quad (\text{Eqn 6.44})$$

If (and only if) L_x and L_y commute, then the above equation may be rewritten as:-

$$\psi^{n+1} = \exp(-L_x \Delta t) \exp(-L_y \Delta t) \psi^n \quad (\text{Eqn 6.45})$$

And this may be further re-arranged into the form:-

$$\exp(L_y \Delta t / 2) \psi^{n+1} = \exp(-L_x \Delta t / 2) \{ \exp(-L_x \Delta t / 2) \exp(-L_y \Delta t / 2) \psi^n \} \quad (\text{Eqn 6.46})$$

since L_x and L_y commute with themselves.

The term in curly brackets can be identified as the half-timestep wavefunction $\psi^{n+1/2}$, and hence equation 6.46 can be written in two parts:

$$\exp(L_y \Delta t / 2) \psi^{n+1} = \exp(-L_x \Delta t / 2) \psi^{n+1/2} \quad (\text{Eqn 6.47})$$

$$\exp(L_x \Delta t / 2) \psi^{n+1/2} = \exp(-L_y \Delta t / 2) \psi^n \quad (\text{Eqn 6.48})$$

Making first-order Taylor expansions of these exponentials (*ie* neglecting terms of order Δt^2 once again) results in the ADI equations of section 6.9.

$$(1 + L_x \Delta t / 2) \psi^{n+1/2} = (1 - L_y \Delta t / 2) \psi^n \quad (\text{Eqn 6.49})$$

$$(1 + L_y \Delta t / 2) \psi^{n+1} = (1 - L_x \Delta t / 2) \psi^{n+1/2} \quad (\text{Eqn 6.50})$$

However, if L_x and L_y do not commute, there will be an error in this splitting process, regardless of how the exponentials are expanded for the purposes of computation .

Taylor expansion of the relevant exponentials gives the following result to first-order in the commutator $[L_x, L_y]$:-

$$\exp(L_x).\exp(L_y) - \exp(L_x+L_y) \approx [L_x, L_y]/2 \quad (\text{Eqn 6.51})$$

Thus, the time-evolution operator $\exp(-L_x\Delta t).\exp(-L_y\Delta t)$ used as the basis of the ADI method in equation 6.45 is not the same as the true composite operator $\exp(-(L_x+L_y)\Delta t)$.

A better approximation to the exact operator would be:-

$$\exp(-L_x\Delta t).\exp(-L_y\Delta t) - 1/2 (-\Delta t)^2 [L_x, L_y] \quad (\text{Eqn 6.52})$$

When the ADI procedure was used to solve the Schrodinger equation for a wavepacket in an AB ring structure, it was found that the stability (as assessed from the normalisation) was good for small magnetic fields, but became poor for magnetic fields above about 0.5T. The instability manifested itself as a gradual but uncontrollable increase in the normalisation.

The spatial discretisation chosen was several times smaller than the cyclotron radius for magnetic field of the order of 1T and the cyclotron energy was also relatively small (at about $1.76\text{meV}\cdot\text{T}^{-1}$) compared to the energy of a typical wavepacket (about 10meV). The conclusion that was drawn from this was that the instability was unlikely to be directly due to the spatial and temporal discretisation errors such as those that will be discussed in chapter 7. Therefore the unitarity of the expansion scheme was investigated.

Denoting the actual operator used in the simulations by O_1 and the exact operator by O_2 , it is seen that O_1 satisfies (approximately) the following identity:-

$$O_1 = O_2 + 1/2 \Delta t^2 [L_x, L_y] \quad (\text{Eqn 6.53})$$

The operator O_1 is tested for unitarity by forming the product $O_1 O_1^+$. To simplify the analysis, only terms to first-order in $[L_x, L_y]$ are retained, so that the unitarity test for O_1 results in the relation:-

$$O_1 O_1^+ \approx 1 + 1/2 \Delta t^2 \{ O_2 [L_x, L_y]^+ + [L_x, L_y] O_2^+ \} \quad (\text{Eqn 6.54})$$

Thus the non-commutivity of the L_x and L_y operators has the effect of making the ADI method a non-unitary process.

A direct calculation of the commutator results in:-

$$\begin{aligned} [L_x, L_y] = & \alpha^2 \{ (m^*/\hbar^2)(d^2V/dy^2 - d^2V/dx^2) - 4i\beta d^2/dxdy \\ & + 1/2 U (d^2/dy^2 - d^2/dx^2) - 1/2 i\beta U (xd/dy + yd/dx) \\ & + i\beta (-1/2 \beta^2 xy - (m^*/\hbar^2)(y dV/dx + xdV/dy)) \\ & + \beta^2 (y(d/dy + xd^2/dxdy) - x(d/dx + yd^2/dydx)) \} \end{aligned} \quad (\text{Eqn 6.55})$$

Where $V=V(x,y)$ is the potential function, $\alpha=-i\hbar/(2m^*)$, $\beta=eB/\hbar$ and $U=U(x,y)$ is a lumped potential energy term, defined as:-

$$U(x,y) = (-e^2 B^2 (x^2 + y^2)/(4\hbar^2) - 2m^* V(x,y)/\hbar^2) \quad (\text{Eqn 6.56})$$

To obtain an estimate of the magnitude of the error term, a typical value for the commutator is estimated, but to make any further progress, some assumptions about the functional form of the potential energy $V(x,y)$ and the wavefunction have to be made. In general there will not be a simple analytic form for $V(x,y)$, and the solution $\psi(x,y)$ is not known in advance, so that the magnitude of each of the terms comprising the commutator cannot in general be found exactly by analytic means. However, a model calculation for a simple case can be performed to obtain an estimate of the order of magnitude of the terms.

Assuming $V(x,y)$ is a constant everywhere in the computational region leads to two-dimensional particle-in-a-box eigenfunctions of the form:-

$$\psi_{k_x, k_y}(x, y) = A \cdot \sin((k_x^2 + V)^{1/2} x) \cdot \sin((k_y^2 + V)^{1/2} y) \quad (\text{Eqn 6.57})$$

Consider first the case for which $V(x,y)$ is a constant and the magnetic field is zero. The commutator in this case reduces to:-

$$[\mathbf{L}_x, \mathbf{L}_y] = \alpha^2 \{ (-m^* V / \hbar^2) (d^2/dy^2 - d^2/dx^2) \} \quad (\text{Eqn 6.58})$$

Evaluating this simplified commutator for the particle-in-a-box eigenfunctions gives an order-of-magnitude estimate for the error term of:-

$$-1/2 \Delta t^2 \alpha^2 m^* V (K_x^2 - K_y^2) / \hbar^2 \quad (\text{Eqn 6.59})$$

Taking the difference in the x and y components of the energy to be around 10meV and a constant potential of 1V (used as the confinement potential in this study) yields a value for the error term of approximately $10^{27} \Delta t^2$. For a timestep of about 10^{-15} s, this term has a relative size of about 10^{-3} .

Consider now the case of a non-zero magnetic field. If the variables x and y are equated to Dx (the mesh size) then for a best-case estimate, a term by term analysis of the commutator shows the last term in β^2 to be the largest. This term is of the order of $\alpha^2 \Delta t^3 10^{37} \text{T}^{-2}$, and has a relative size (for a timestep of 10^{-15} s) of about 1.0T^{-2} ; a factor of about 1000 greater (per Tesla^2) than the zero-field error term.

Therefore the application of a magnetic field of the order of 1 Tesla can destroy the unitarity of the ADI expansion scheme. As a result, the simulations in this study were limited to fields in the order of about 10mT. It is possible that the stability of the ADI scheme could be improved for high-field calculations by taking the higher-order corrections involving the commutator $[\mathbf{L}_x, \mathbf{L}_y]$ into account.

6.11 Vector-processing techniques employed for high-speed computation.

In late 1988 the University of Glasgow acquired an IBM 3090 model 150E/VF mainframe computer. This project was fortunate enough to be chosen as one of several special study projects which benefited from a close collaboration with IBM under the Kelvin Project, the purpose of which was to assess the performance of numerically intensive applications on the IBM 3090 in a university environment.

Since the IBM 3090 is capable of processing in vector mode, the CPU can work on several pairs of numbers simultaneously in its internal registers. For example, a multiplication of two floating-point numbers might typically consist of four distinct stages:-

- 1) Addition of exponents
- 2) Pre-normalisation exponent adjustment
- 3) Multiplication of mantissas
- 4) Post-normalisation exponent adjustment

In normal, or scalar operation, a particular pair of numbers must complete all four steps before the next pair can be processed. In vector mode however, the processing of a second pair of numbers can be started as soon as the first pair moves onto stage two of their processing. A third pair can then be processed after the first pair has moved to stage three and the second pair to stage two, and so on. The CPU can be thus working on four pairs of numbers simultaneously. Vector processing can offer significant increases in speed over scalar operation in computations involving repetitive operations on large arrays.

6.11.1 Modification of computer code for vector processing.

Considering the first half-sweep of the mesh in the ADI procedure, the tri-diagonal algorithm described in section 6.8 generates FORTRAN code with the following structure:-

```
DO 10 IY=1, IYMAX ! For each row
C      Forward elimination stage.
      {Initialise values for start of recursion formula for IX=1}
      DO 20 IX=2, ( IXMAX -1 )
          W(IX) = Fn{ W(IX-1) }
          G(IX) = Fn{ W(IX-1), G(IX-1) }
20      CONTINUE
      { Fix end-values of recursion formula for IX= IXMAX }
C      Back-substitution stage.
      DO 30 IX=(IXMAX -1 ), 1, -1
          PSI( IX,IY ) = Fn{ PSI(IX+1,IY), W(IX), G(IX) }
30      CONTINUE
10      CONTINUE
```

Where the notation $F_n\{ \}$ denotes the functional dependence of the variable on the left hand side. The DO 20 and DO 30 loops are recursive, which means that the result from the previous computation must be known in full before the next pair of numbers can be started in the CPU vector pipeline. Vectorisation of this code as it stands therefore, is impossible.

The only other possibility is to vectorise on the DO 10 loop. The compiler will not do this automatically because on each cycle through the DO 10 loop, the values of $W(IX)$ and $G(IX)$ will be different. To overcome this, extra dedicated storage has to be allocated, as the compiler does not provide for the allocation of any storage dynamically during execution.

Therefore, the code was modified so that it had the following structure:-

```

DO 10 IY=1, IYMAX ! For each row
C      Forward elimination stage.
      {Initialise values for start of recursion formula for IX=1}
          DO 20 IX=2, ( IXMAX -1 )
              W(IX,IY) = Fn{ W(IX-1, IY) }
              G(IX,IY) = Fn{ W(IX-1, IY), G(IX-1, IY) }
20          CONTINUE
      { Fix end-values of recursion formula for IX= IXMAX }
C      Back-substitution stage.
          DO 30 IX=(IXMAX -1 ), 1, -1
              PSI(IX,IY)=Fn{PSI(IX+1,IY), W(IX,IY), G(IX,IY)}
30          CONTINUE
10      CONTINUE

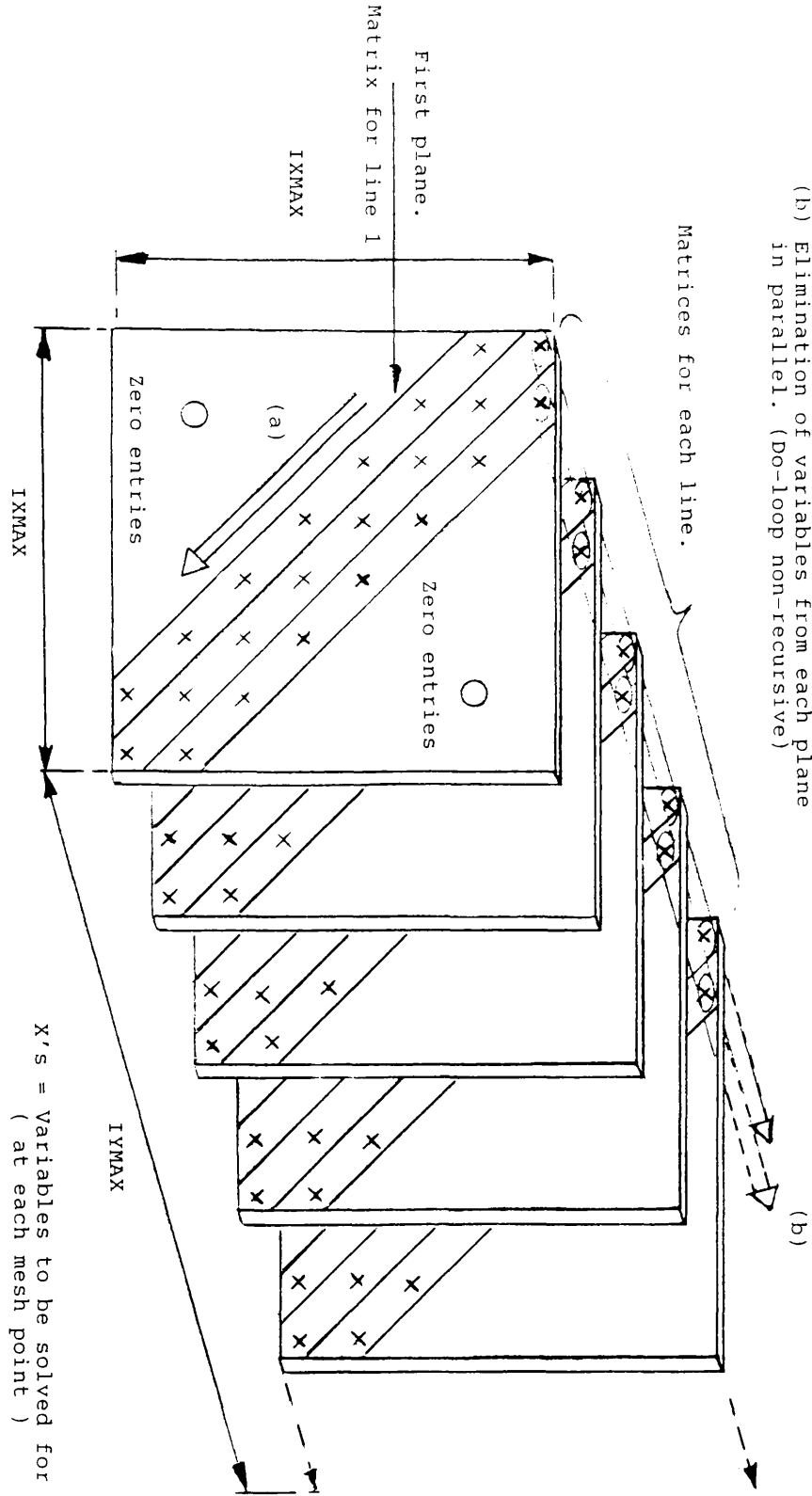
```

Figure 6.1 shows how this process may be visualised. The calculation for any particular half-sweep can be thought of as solving many sets of *independent* one-dimensional problems. For the code shown, the wavefunction for each row is found as a function of IX and is independent of all other rows (in the y-direction).

Since vector processing was found to be impossible for each matrix in the x-direction (referring to figure 6.1), the variables can be thought of as being stored in a cubic format and vectorisation performed in the y-direction for which there are no recurrences. The disadvantage of this technique is that more storage is required and therefore the range of the IY-index on which vectorisation was performed was limited to 128 at a time so that in this case the complete half-sweep computation consisted of two such cubes. The length of 128 variables was chosen because this was the number of variables the internal CPU registers could hold at a time. Longer vector lengths result in faster processing speeds, but for vectors longer than 128, the increase in speed with length becomes a law of diminishing returns.

Figure 6.1. Allocation of extra storage to break recursive data dependences.

- (a) Usual direction of elimination - repeated for each plane (Do-loop recursive).
- (b) Elimination of variables from each plane in parallel. (Do-loop non-recursive)



The second type of modification concerns the efficient use of the high-speed cache memory. The cache memory holds 64 kBytes of data from the main memory which the CPU can then access as fast as necessary without incurring the relatively long data fetch time from the main memory.

There is however a problem that sometimes arises when two-dimensional arrays are processed. In FORTRAN, 2D arrays are physically stored in a one-dimensional format in "column major" order so that, for instance, the 3 by 3 array

1	2	3
4	5	6
7	8	9

would be internally stored in consecutive memory locations as:-

1 4 7 2 5 8 3 6 9

Thus if the array is accessed in column major order with the left hand subscript varying most rapidly, the internal memory locations are addressed in order, one after another. On the other hand, if the array is addressed in row major order, the internal memory locations are accessed every 3 elements. The distance between consecutive array elements as they are accessed is called the stride. For an N by N array, row major addressing results in a stride of N, whereas column major addressing always results in a stride of 1.

The high-speed cache is organised in lines of 128 bytes, which are loaded one line at a time from the main memory as required. Processing with a stride of 1 is the most efficient mode of operation because each cache line is fully accessed before it is necessary to load another line. However, the larger the stride, the more frequently the cache lines must be loaded and this increases the average data fetch time by increasing the number of references to the main memory.

The vectorised code was therefore written as follows, using the array G(IY,IX) as a temporary store for a transposed table of PSI(IX,IY).

```

DO 10 IY=1, IYMAX ! For each row
C   Forward elimination stage.
    {Initialise values for start of recursion formula for IX=1}
        DO 20 IX=2, ( IXMAX -1 )
            W(IY,IX) = Fn{ W(IY, IX-1) }
            G(IY,IX) = Fn{ W(IY,IX-1), G(IY,IX-1) }
20    CONTINUE
    { Fix end-values of recursion formula for IX= IXMAX }
C   Back-substitution stage.
        DO 30 IX=(IXMAX -1 ), 1, -1
            G(IY,IX)=Fn{ G(IY,IX+1), W(IY,IX), G(IY,IX)}
30    CONTINUE
10    CONTINUE

```

This enables stride-one processing in the "direction of vectorisation" on the IY variable.

To recover the array PSI(IX,IY), a separate loop, outside the main IY-vectorised loop, was inserted to transpose the array G(IY,IX).

```

DO 40 IX = 1, IXMAX
    DO 40 IY = 1, IYMAX
        PSI(IX,IY) = G(IY,IX)
40    CONTINUE

```

This loop was then vectorised on the IX variable.

Other smaller modifications such as the use of compound multiply-and-add instructions and of an optimised library routine for the calculation of a scalar-product were also made.

The modification of the original code so that it was in a form suitable for vector-processing enabled almost all of the computation time to be spent in vector mode compared to almost none using the unmodified code. This resulted in about a five-fold increase in computation speed over normal scalar computation.

6.12 Summary.

This chapter has discussed the general aspects of a two-dimensional quantum ballistic transport model.

The efficient solution of the two-dimensional problem is non-trivial, and it is shown that schemes based on a finite expansion of the time evolution operator, which evolve the solution on a timestep basis, offer substantial savings in computational effort over more traditional eigenfunction expansion techniques.

The expansion of the time-evolution operator must preserve the properties of the original equation as far as possible. In particular it must result in a stable numerical scheme which also preserves unitarity. The Cayley form of the time-evolution operator, which arose from consideration of the Crank-Nicholson scheme, was shown to meet both these criteria.

Although one could tackle the two-dimensional problem by expansion of the time-evolution operator in terms of the full Hamiltonian matrix, it was shown that by resolving the Hamiltonian into two components, a solution could be obtained by multiple one-dimensional solutions along the rows and columns of the mesh. The advantage of doing this is that the very-efficient tri-diagonal algorithm described in section 6.8 can be used.

Despite the use of these techniques, the computation was still very numerically intensive and attention was therefore given to the optimum use of the vector-processing facility on the IBM3090. The original code written was impossible to vectorise on the IBM due to the recursions in the resulting formulae. However, since all the separate 1D problems are independent for each half-timestep, it was shown how the code for the full 2D calculation could be vectorised by the allocation of extra storage.

7. Computer modelling of an Aharonov-Bohm ring.

7.1 Introduction.

This chapter describes the details of a two-dimensional computer model for an AB ring.

The Hamiltonians for two different magnetic field distributions are derived and it is shown how these are resolved into two components necessary for the solution of the numerical problem.

Although the computation scheme chosen was shown to be both stable and unitary in chapter 6, this does not necessarily mean that it is accurate. To assess the accuracy of the model, the spatial and temporal discretisation is considered. These considerations in addition to those imposed by the stability requirement lead to the final choice for the mesh size and timestep.

The choice of the initial condition is also discussed and a technique for modelling the contacts at the mesh boundaries is described. Successful modelling of the contact regions was of great importance in reducing the computational effort required for subsequent magneto-transmission calculations.

7.2 Construction of the Hamiltonian.

For motion in a magnetic field, the one-electron Hamiltonian takes the following form:-

$$H = (\mathbf{P} - e\mathbf{A})^2/2m^* + V(x,y) \quad (\text{Eqn 7.1})$$

where \mathbf{A} is the vector potential arising from the magnetic field. Spin has been neglected because at all but the highest fields, the contribution of the spin term is comparatively small and the inclusion of spin would also complicate matters by the need to represent the wavefunction as a two-component spinor.

If the magnetic field is considered to be perpendicular to the plane of the ring (in the z-direction) and the symmetric gauge is used for the vector potential, then:-

$$\mathbf{A} = 1/2 B (-y, x, 0) \quad (\text{for } \mathbf{B} = (0, 0, B)) \quad (\text{Eqn 7.2})$$

The Hamiltonian may then be expanded to give:-

$$\begin{aligned} H = & (\hbar^2/2m^*)(\nabla^2 - 2m^*V(x,y)/\hbar^2 - e^2B^2(x^2+y^2)/(4\hbar^2) \\ & + (-i\hbar)eB(x.d/dy - y.d/dx)/\hbar^2) \end{aligned} \quad (\text{Eqn 7.3})$$

where use has been made of the commutators

$$[y, p_x] = 0 \quad \text{and} \quad [x, p_y] = 0 \quad (\text{Eqns 7.4})$$

Rewriting the Schrodinger equation in a simplified form:-

$$d\psi/dt = -L \psi \quad (\text{Eqn 7.5})$$

gives

$$\begin{aligned} L = & (-i\hbar/2m^*)(\nabla^2 - 2m^*V(x,y)/\hbar^2 - e^2B^2(x^2+y^2)/(4\hbar^2) \\ & + ieB(y.d/dx - x.d/dy)/\hbar) \end{aligned} \quad (\text{Eqn 7.6})$$

Two components of L may then be identified as:-

$$L_x = \alpha(d^2/dx^2 + 1/2 \cdot U(x,y) + i\beta y.d/dx) \quad (\text{Eqn 7.7})$$

$$L_y = \alpha(d^2/dy^2 + 1/2 \cdot U(x,y) - i\beta x.d/dy) \quad (\text{Eqn 7.8})$$

$$\text{where } \alpha = -i\hbar/(2m^*), \quad \beta = eB/\hbar \quad (\text{Eqns 7.9})$$

and

$$U(x,y) = -e^2B^2(x^2+y^2)/(4\hbar^2) - 2m^*V(x,y)/\hbar^2 \quad (\text{Eqn 7.10})$$

entering as a lumped potential energy term which has been split equally between the two operators.

Using the ADI procedure described in the last section, the equations to be solved are

$$(1 + \Delta t/2 L_x) \psi^{n+1/2} = (1 - \Delta t/2 L_y) \psi^n \quad (\text{Eqn 7.11})$$

$$(1 + \Delta t/2 L_y) \psi^{n+1} = (1 - \Delta t/2 L_x) \psi^{n+1/2} \quad (\text{Eqn 7.12})$$

Each of these equations is then solved in turn using the tri-diagonal algorithm described in section 6.8.

7.3 Construction of the Hamiltonian for the case of a magnetic flux piercing the centre of an AB ring

In this case, it is assumed that the magnetic field distribution is entirely within the ring so that electrons in the ring experience a vector potential but no force through the magnetic field.

For the case when the magnetic field is assumed to be in the z-direction, initial working is best done in polar co-ordinates. Starting with

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (\text{Eqn 7.13})$$

as the definition of the vector potential, surface integration of the magnetic field over the interior of the ring results in the relation:-

$$\oint \mathbf{A} \cdot d\mathbf{l} = \int \mathbf{B} \cdot d\mathbf{s} \quad (\text{Eqn 7.14})$$

For a single line of magnetic flux or "magnetic string" the A_z and A_r components of the vector potential are zero, so that the magnetic field in the z-direction only gives rise to a rotational component, A_θ . From symmetry considerations A_θ must be a constant for a fixed radius away from the centre.

Performing the line integral in equation 7.14 for a fixed radius gives:-

$$\oint \mathbf{A} \, d\mathbf{l} = A_\theta \, 2\pi r \quad (\text{Eqn 7.15})$$

But $\oint \mathbf{B} \, d\mathbf{s}$ is just the total amount of flux threading the ring Φ , so that the theta component of the vector potential is:-

$$A_\theta = \Phi / 2\pi r \quad (\text{Eqn 7.16})$$

In numerical work however, rectangular polar co-ordinates are generally easier to work with, so that the A_θ component of the vector potential is expressed in terms of x and y components for the purposes of computation.

Expansion of the Hamiltonian in equation 7.1 proceeds as before for the case of the uniform magnetic field, except in this instance the components of momentum and vector potential do not commute with each other. Specifically, the relations are:-

$$[p_y, A_y] = (-i\hbar/2\pi) \Phi \left(-2xy/(x^2+y^2)^2 \right) \quad (\text{Eqn 7.17})$$

$$[p_x, A_x] = (-i\hbar/2\pi) \Phi \left(+2xy/(x^2+y^2)^2 \right) \quad (\text{Eqn 7.18})$$

Fortunately, it is seen that these commutators are equal and opposite in sign, so that a reasonably tractable form for the Hamiltonian is still recovered:-

$$\begin{aligned} H = & (-\hbar^2/2m^*)(\nabla^2 - e^2\Phi_b^2/(\hbar^2(x^2+y^2))) \\ & -ie\Phi_b.(2x \, d/dy - 2y \, d/dx) /(\hbar^2(x^2+y^2)) - 2m^* V(x,y)/\hbar^2 \end{aligned}$$

$$(\Phi_b = \Phi/2\pi) \quad (\text{Eqn 7.19})$$

The magnetic string was considered to be placed at the centre of the ring so that in the actual computer model, an additional transformation was performed to re-locate the origin at the centre of the ring instead of at the bottom left-hand corner of the region.

7.4 Specification of the initial condition.

The initial condition is represented by a plane wave of wavevector k_0 , multiplied by a Gaussian envelope. The width of the wavepacket represents the initial uncertainty of the electron position and the wavevector k_0 specifies the average forward momentum. The initial wavefunction in one-dimension has the form:-

$$\psi(x) = A \exp(-x^2/(2\sigma^2)) \cdot \exp(ik_0x) \quad (\text{Eqn 7.20})$$

The y-variation is derived from the form of the lowest eigenfunction for an infinite well using the wire width, L_w , as the well width yielding a full two-dimensional initial wavefunction of the form:-

$$\psi(x,y) = A \exp(-x^2/(2\sigma^2) + ik_0x) \cdot \cos(\pi(y-y_0)/L_w) \quad (\text{Eqn 7.21})$$

When specifying the initial distribution, one must ensure that the Gaussian envelope is not too severely truncated by the boundary of the computational grid. The errors involved due to truncation of the initial condition will be discussed in detail in section 7.5, but for the moment a few general comments are made about the size of the wavepacket.

If a free wavepacket is injected from a contact far away from the region under study and is allowed to travel ballistically, the width of the wavepacket becomes broader and the corresponding Fourier spectrum becomes sharper. For instance, a 10meV wavepacket with an initial standard deviation of 50nm doubles in width roughly every half micron. This might lead one to suspect that a realistic form for the wavepacket which is some way away (several microns) from the contact would be a very broad, plane-wave like, Gaussian. In reality, the wavepacket does not keep broadening indefinitely due to the fact that the electron scatters inelastically from time to time. The process of inelastic scattering can be thought of as a measurement process on the electron and thus serves to identify the position of the wavepacket to within some region.

An argument put forward by Thouless [1980] treats the wavepacket in this manner and the diffusion constant D is calculated on the basis that each time the wavepacket

suffers an inelastic collision it finds itself in some confining region of length l .

For long inelastic scattering times τ_i , the diffusion constant was found to be:-

$$D = l^2/12\tau_i \quad (\text{Eqn 7.22})$$

For the purposes of this study, a detailed knowledge of the width of the wavepacket is probably unimportant, but it is worthwhile to note that the extent of the wavepacket is indeed limited, and the confining length is proportional to the square root of the inelastic scattering time.

The view taken in this study was accept the presence of inelastic scattering outside the region of interest (giving rise to a limited-breadth wavepacket), but to consider the transport to be entirely ballistic within the region of interest. If quantum effects are to be seen at all, the wavepacket must be free from inelastic collisions during its passage through the device. Thus, it would seem reasonable to take a standard deviation for the wavepacket which is at least about the same order as the ring diameter; the actual value being chosen to avoid large truncation errors in the limited space available on the computational grid.

In preliminary work, the energy was chosen such that the DeBroglie wavelength was commensurate with the dimensions of the ring. Thus for a test ring of 300 Angstroms diameter, an energy of 0.15eV was chosen (λ approximately 120 Angstroms). For subsequent larger-scale simulations, the energy used was the Fermi energy deduced from the sheet electron concentrations measured in actual devices (Appendix C).

7.5 Errors incurred in the process of discretisation.

In this section the various criteria which must be met in choosing physical parameters such as spatial mesh size and timestep size are examined.

Goldberg *et al* [1967] have analysed the discretisation of a similar problem in one dimension and analysis of the errors caused by the spatial differencing in a null potential region resulted in the constraint:-

$$K_m^2 \Delta x^2 / 12 \ll 1 \quad (\text{Eqn 7.23})$$

where K_m is the wavevector of the highest Fourier component in the system.

This estimate arose from a consideration of how accurately the finite difference form approximated the known eigenfunctions of the "particle-in-a-box" problem. Introducing a scalar potential of V_{max} everywhere, added to the kinetic energy term, changing the above constraint to:-

$$(K_m^2 + V_{max}) \Delta x^2 / 12 \ll 1 \quad (\text{Eqn 7.24})$$

The highest wavevector of interest can be found by Fourier transformation of the initial spatial distribution of the wavepacket. Writing the momentum distribution in a form similar to the spatial distribution:-

$$\psi(k) = A \exp(ikx_0) \exp(-(k-k_0)^2 / (2\sigma_k^2)) \quad (\text{Eqn 7.25})$$

Upon which the identity $\sigma_k = 1/\sigma_x$ is established.

The energy chosen for the test simulation was 0.15 electron volts, which gave a wavepacket with an average wavevector of about $130 \times 10^6 \text{ m}^{-1}$. The spatial standard deviation was chosen to be 10 times the mesh cell size, as this was about the maximum size the input wire could contain due to memory constraints on the computer in use at the time. The table below shows the values of the wavefunction in momentum or real space for various values of position or wavevector away from the central average (normalised to unity).

Table 7.1

<u>$(k-k_0)$ or $(x-x_0)$</u>	<u>$\psi(x)$ or $\psi(k)$</u>
σ	0.6
2σ	0.1
3σ	0.01

This shows that if the Gaussian wavepacket is truncated at a wavevector of between two and three times the standard momentum deviation above the average, then this can be considered an acceptable truncation error.

Taking the maximum wavevector of interest to be

$$K_{\max} = K_{\text{ave}} + 3\sigma_k \quad (\text{Eqn } 7.26)$$

and substituting this value into the stability requirement of equation(7.23), suggests a mesh cell size of less than about 3×10^{-9} m. This value ensures that all wavevector values up to K_{\max} are faithfully represented on the discretised mesh.

The Gaussian is also truncated in real space by the boundaries of the model. At the boundaries, the wavefunction is held at zero, which is different to the proper value if the Gaussian were infinite in extent. From table 7.1, it is seen that an acceptable truncation error would again be between two and three standard deviations from the centre of the wavepacket. In the test simulation, the wavepacket was about $2.7\sigma_x$ away from the left hand side mesh boundary, giving a truncation error of about 1%. Unfortunately, because the simulation was performed on a two-dimensional mesh, it was not possible to be as generous as Goldberg *et al* [1967] in the allocation of mesh points to avoid truncation errors. This was especially noticeable when the wavepacket was propagated for any reasonable length of time, when the transmitted and reflected portions of the wavepacket interacted with the mesh boundaries. In one dimension Goldberg *et al* were able to make their computational region large enough so that negligible interaction with the boundaries took place during the simulation without incurring unreasonably long computations. Hence, the criteria of Goldberg associated with this requirement namely that of ensuring sufficient space for the wavepacket during the simulation are not satisfied.

7.6 Choice of an appropriate timestep.

Having chosen the size of ring for the test simulation, the available machine memory allowed a mesh cell size of 5 Angstroms to be taken, which is within the stability requirement demanded by equation 7.23.

To obtain an estimate for a suitable timestep, the Schrodinger equation can be compared to the closely-related diffusion equation for which the stability conditions have been extensively studied. Treating the Schrodinger equation as an imaginary-time diffusion equation, one can write:-

$$D \frac{d^2\psi}{dx^2} = d\psi/d\tau, \quad \tau=it, \quad D = \hbar^2/2m^* \quad (\text{Eqn 7.27})$$

where D is the diffusion constant and τ is an imaginary-time parameter. A suitable timestep may be estimated by applying the Courant-Friedrichs-Lewy criterion [Potter1973]. This states that the maximum velocity that can be represented by the temporal and spatial differencing, $\Delta x/\Delta t$, must be at least as large as the maximum actual velocities in the physical system.

$$|V_{\text{numerical}}| \geq |V_{\text{physical}}| \quad (\text{Eqn 7.28})$$

In the simulation, wavevectors larger than $\pi/\Delta x$ cannot be represented by the spatial differencing method in any case, so that the maximum physical velocity of interest is the group velocity associated with wavevector $\pi/\Delta x$. This results in the requirement:-

$$\Delta x^2/\Delta t \geq 2\pi D, \quad D = \hbar^2/2m^* \quad (\text{Eqns 7.29})$$

Thus, the smaller the mesh cell size, the smaller the time interval must be for sampling. The Δx^2 dependence can be seen to arise from the diffusion length over a time Δt , through the classical diffusion relation:-

$$\Delta x = (D \Delta t)^{1/2} \quad (\text{Eqn 7.30})$$

This requirement suggests a maximum timestep size of about 1.5×10^{-16} second for the energy used.

7.7 Errors in the phase of the wavefunction.

In addition to the errors incurred by the spatial differencing of the continuous space, the finite time discretisation also leads to errors. The Courant-Friedrichs-Lewy criterion concerns itself with the stability of the solution and must always be observed. Having thus placed an upper limit on the timestep size, this section obtains estimates of the errors caused by the temporal differencing.

Following Goldberg *et al*, the null potential is again considered and the Cayley approximation for the time-evolution operator is applied to a plane wave and then compared to the exact expression. The result for the maximum *relative* error in the Fourier components is (in SI units):-

$$\text{Error} = (N\Delta t^3/12) (\hbar/2m^*)^3 (K_m^6 - K_0^6) \ll 1 \quad (\text{Eqn 7.31})$$

Where the wavevector K_m is the highest Fourier component and K_0 the average. The number of timesteps is N and Δt is the timestep size.

It was estimated that the wavepacket would take in the order of 1×10^{-13} sec. to propagate around the ring, and a few trial computations suggested a suitable total simulation time of about 1.5×10^{-13} sec. The timestep was then chosen so that the relative phase error could be kept below 1% or less during this period of time. The final choice for the timestep was 2×10^{-17} second over 8000 ^{timesteps} giving a maximum relative phase error in the test simulation of about 6.5×10^{-3} .

7.8 Modelling the contacts.

It was intended to use the computer model to measure the total transmission and reflection coefficients of the ring, that is the total proportions of reflected and transmitted wavepackets after all interactions with the ring are complete. This task posed a problem of how best to model the boundaries or "contacts" in the simulation. The wavefunctions in figures 8.2 to 8.4 show standing wave structure at the output and input wires in the direction of motion. This is an artifact of the numerical procedure caused by demanding that $\psi(x,y)=0$ at the boundaries, and is not physical. Unfortunately the standing waves, caused by the transmitted and reflected waves rebounding off the boundaries, made the determination of transmission coefficients

very inaccurate.

A possible solution would have been to extend the length of the modelled system so that by the time the reflected and transmitted wavepackets reach the boundaries, interaction with the ring is virtually complete. The problem with this strategy is that the length of the system would have to be made unreasonably large, causing an unacceptable increase in an already lengthy computation.

The second method is to consider how the wavefunction can be absorbed by modelling a contact. Ideally, such a contact should totally absorb any wavefunction impinging on it whilst reflecting nothing. In practice such a contact could not be made totally absorbing or without a certain amount of reflection. However, by increasing the length of the contact, both these parameters could be made arbitrarily small. Furthermore, the resulting length of the contact was much less than what would have been required if no absorption had occurred.

To absorb the wavefunction, an imaginary part to the potential was introduced. A proof that a wavefunction in a complex potential is completely absorbed is given below.

Consider a model Hamiltonian of the form:-

$$H = (-\hbar^2/(2m^*)) \nabla^2 + (V_r + iV_i) \quad (\text{Eqn 7.32})$$

where V_r is the real part of the potential and V_i the imaginary part. This operator is non-Hermitian and so when used in the time-evolution operator will not conserve the normalisation of the wavefunction. For an absorber it is required that the normalisation decrease, corresponding to a loss of particles into the contact. The general result for the normalisation as a function of time is:-

$$\text{Norm}(t) = \int \psi(r,t) \psi^*(r,t) d^3r \quad (\text{Eqn 7.33})$$

One can now substitute for $\psi(r,t)$ from the time-dependent Schrodinger equation using the model Hamiltonian to yield,

$$\begin{aligned} \text{Norm}(t) = & \int \exp(-it/\hbar) (-\hbar^2/(2m^*)) \nabla^2 + (V_r + iV_i) \psi(r,0) \\ & \exp(+it/\hbar) (-\hbar^2/(2m^*)) \nabla^2 + (V_r - iV_i) \psi^*(r,0) d^3r \end{aligned}$$

$$\begin{aligned}
&= \int \exp(-2V_i t / \hbar) \psi(r,0) \psi^*(r,0) d^3r \\
&= \exp(-2V_i t / \hbar) \text{Norm}(0)
\end{aligned}
\tag{Eqn 7.34}$$

The normalisation thus decreases exponentially with time if the imaginary part of the potential is made negative. In the simulation the wavefunction is not always situated in a complex potential, but moves into it from a region of zero potential. Reflection at the contact arises in two ways; that caused by the potential discontinuity (similar to real-potential barriers) and that caused by reflection of the (damped) wave reflecting off the boundary and back out of the contact. The initial reflection when the wave first impinges on the contact is proportional to the height of the imaginary potential, and so the amount of reflection can be reduced by lowering the imaginary potential. However, if the potential is lowered too much, the wave takes longer to become damped and can then rebound off the boundary and back out of the contact without sufficient attenuation. The latter effect can be reduced by increasing the length of the contact, causing the wave to dwell longer in the damping region. This does of course increase the amount of computation and therefore a balance must be reached between the amount of acceptable reflection and the length of the contact region. The reflection can be reduced further by smoothly grading the imaginary potential profile so that sharp discontinuities in potential are avoided and in this work a Gaussian profile was chosen. Reasonable parameters for the contact were found empirically using a one-dimensional simulation to study the behaviour of a wavepacket incident on such a potential. It was assumed that a good range of values for the peak complex potential would be roughly the same as the wavepacket energy. From this, the length of the contact was adjusted so that an acceptable reflectance was obtained. An acceptable reflectance was considered to be less than about 1% because this was about the same order as the other errors in the simulation. The final choice was a peak potential of 50mV and a length of 830 Angstroms.

There is a considerable advantage of using this technique over simply increasing the size of the grid. Using a mesh cell size of 16.6 Angstroms and a wavepacket energy of 10meV, the wavepacket took about 50ps to fully interact with the ring in a region of 512x16.6 Angstroms length. The interaction was judged to be complete when the wavefunction normalisation representing that part of the wavepacket still trapped inside the ring fell below 1% of the initial value (1.00).

The average (ballistic) velocity of the 10meV wavepacket in GaAs is about 230×10^3 m/sec, which means that a region at least 11.5 microns would be required to contain the wavepacket for the necessary 50ps. In actual fact, this probably under-estimates the space necessary, because the wavepacket generally splits into transmitted and reflected components on impacting the ring. Thus the space required is nearer twice the original estimate. Even then, the high-velocity components of the wavepacket have been ignored. A region approximately 23 microns long is therefore required to model a region 0.85 of a micron; an increase of about 2600%. In contrast, the contacts modelled using the complex potential occupy 0.083 micron at each end, adding only 20% to the total length.

7.9 Summary.

This chapter has considered the specific aspects of a two-dimensional model for an AB ring. The Hamiltonians for two different magnetic field distributions were derived and it was shown how these were resolved into the two components necessary to implement the ADI method which was described in general terms in chapter 6. The particular decomposition used was not unique, but partitioning the lumped potential energy term equally between the two operators was considered sensible.

The spatial and temporal discretisation errors were estimated by considering how accurately the finite-difference equations approximated the known eigenfunctions for a null-potential region. For the preliminary calculations on the small ring, the error due to the truncation of the initial condition was estimated to be about 1% with an error in the phase of the wavefunction of about 0.5% and a negligible error due the spatial differencing. Errors for the full-scale calculations were slightly worse, with a relative phase error of about 2% and a error due to spatial differencing of around 1%. However, in view of the other uncertainties encountered in constructing the model, such as determining the correct wavepacket energy and the precise form of the confinement potential, these errors were considered acceptable.

A method of modelling the contacts of the device by the addition of an imaginary potential to the Hamiltonian was described. The trade-off between the width and absorbing efficiency of the contact was discussed and it was found that acceptable results could be obtained by a 0.166 micron increase in length of the modelled region. This compares to about an extra 23 microns if the length of the region were simply extended to avoid interference occurring at the boundaries.

8. Two-dimensional simulation results.

8.1 Introduction.

In this chapter the computer model described in the previous chapter is used to simulate the propagation of a wavepacket through an AB ring.

In the first instance a small-scale simulation was performed, which demonstrated the AB effect in a two-dimensional system, although unphysical boundary effects limited the length of time the simulation could be run.

Subsequently, larger, more realistically sized rings were simulated in which the boundaries were modified to remove most of the unphysical boundary reflections. The results show a much more complex distribution of the wavefunction than found in the preliminary studies and the emergence of a multi-mode structure in the wavefunction across the wires is observed.

The non-ideal magneto-transmission characteristic for the larger ring, which is dealt with in greater detail in the next chapter, is discussed in terms of the distribution of the interference fringes with respect to the output wire.

The transit time of the wavepacket through the ring is also considered and was approached in two ways. By measuring the position of the wavefront and by measuring the transmitted charge as a function of time. Both results were compared to the analytic result for a free wavepacket moving in one dimension.

8.2 Description of the idealised AB ring modelled.

Figure 8.1 shows a plan view of the idealised patterned-gate HEMT structure used for the computer simulations. The shaded regions denote areas of high scalar potential of 100 times the average wavepacket energy. The areas not shaded are regions of zero scalar potential which represent the wires forming the ring. The wavefunction is contained for the most part in the wire regions, but since the confining walls are not infinite in height, some penetration into the high-potential regions does occur. The potential profile across the wire was taken to be a perfectly rectangular well. A calculation by Davies [1988] has shown that in the limit of very few carriers in the the wires, the potential was of parabolic shape, however, a self-consistent solution of

the Schrodinger and Poisson equation for a wire with a higher carrier density [Laux *et al* 1988] showed that the potential distribution flattened-out the bottom of the parabolic potential, making it more rectangular. This being so, the original assumption of a rectangular potential across the wire was considered to be of acceptable accuracy. The non self-consistent calculations extended to the case of the wires forming a ring were qualitatively the same [Davies 1988]. The effect of the formation of the wires into a ring was to distort the parabola by a small amount so that the minimum in potential occurred slightly further towards the centre of the ring than at the geometric centre of the wire.

8.3 Preliminary results.

Preliminary calculations were performed for an AB ring of 300Angstroms outside diameter using an HP9000 desktop computer. The wire width was taken to be 50Angstroms and a test energy of 0.15eV was taken. The diameter of 300 Angstroms taken for these preliminary studies was much less than typical fabricated diameters of 1-2 microns because of restrictions on computer memory limiting the number on grid points that could be used. Nevertheless, it was hoped that this model would provide some results on the effects of the two-dimensionality of the ring and give confidence for attempting a more realistic calculation.

Figures 8.2 to 8.4 show a time-sequence of frames showing the propagation of a Gaussian wavepacket (initial standard deviation 50Angstroms) through an AB ring. The photographs show the distribution of the position probability density $|\psi(x,y)|^2$, the value of which is indicated on the linear colour scale (increasing from left to right) shown above each set of frames. The probability density was represented by a scale of 15 colours (plus a black background) so that any areas with a density less than 1/16th of the maximum value were left black. It was found on a trial and error basis that the most satisfactory visual representation was to scale the distribution to the maximum density at each frame rather than from the maximum density in the initial condition. This was because it was of more interest to observe the general distribution of the probability distribution rather than the absolute value, especially in view of the fact that the wavepacket quickly delocalised on impact with the ring.

The simulations performed without any applied magnetic field show that the wavepacket splits symmetrically at the input junction with about 30% of the wavepacket being reflected back along the input wire. The two partial wavepackets in each arm of the ring then go on to interfere constructively with each other leading to a peak in the probability density centred on the output wire. This peak was observed to propagate out of the ring and along the output wire. The standing-wave structure in the reflected and transmitted wavepacket is due to reflection from the boundary of the computational mesh, which acted as an infinite potential due to the boundary condition $\psi(\text{boundary})=0$. This was seen more clearly in the larger simulations in which the boundaries were further away from the exit of the ring. The standing-wave structure seen in the emerging wavepacket in the smaller simulations therefore is not due to any special property of the ring itself.

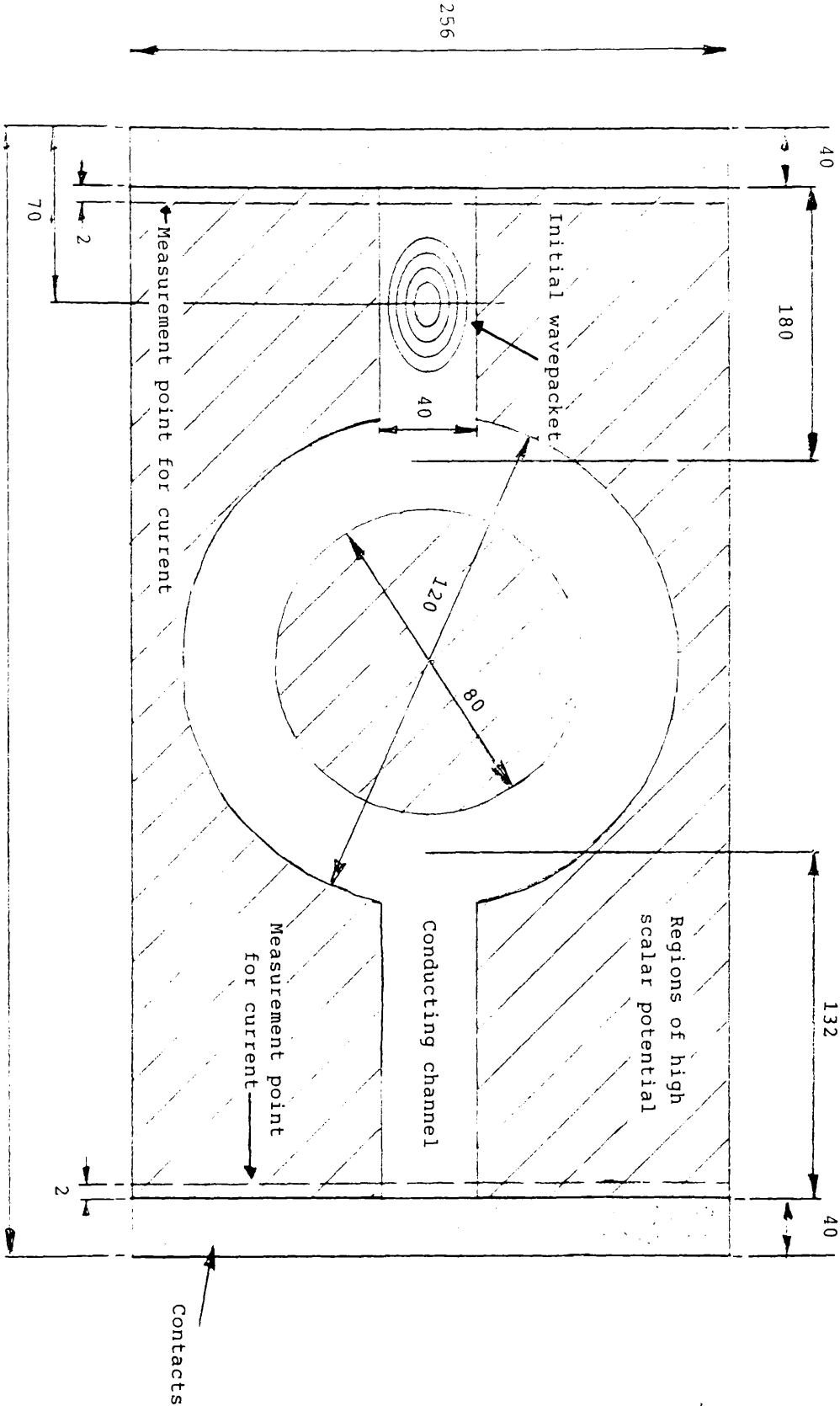
A major problem in the preliminary calculations was that even if the unphysical boundary reflections were initially ignored, they limited the length of time the simulation could be run, as they ultimately began to interfere with the wavefunction throughout the entire region.

Figures 8.5 to 8.7 show a sequence of frames from a simulation of the same ring when a magnetic string of strength $0.5h/e$ was positioned in the centre of the ring. Initially, the same behaviour was noted, but this time the square-modulus of the wavefunction had a minimum value centred on the output wire and resulted in negligible transmission of the wavefunction out of the ring.

Figures 8.8 and 8.9 show a three-dimensional representation of the wavefunction at 46fs from launch in the absence and presence of a magnetic string respectively. The viewpoint of these plots is from the direction of the output wire, so that the large peak at the top of both plots is the reflected wavepacket and the other peaks are the interference fringes at the output junction. These two plots further clarify the role of the magnetic flux in changing the central peak in the interference pattern into a central trough.

These preliminary results directly illustrate the theoretical existence of the AB effect in a truly two-dimensional system. In practice however, the applied magnetic field is always uniform (not string-like) over the very small area the ring occupies. The sequence in figures 8.10 to 8.12 show the results of a simulation of this situation.

Figure 8.1 An idealised Aharonov-Bohm ring structure used in computer modelling.



Figures 8.2 to 8.4: Propagation of a Gaussian wavepacket through a 30nm diameter test ring with no applied field.

Fig 8.2

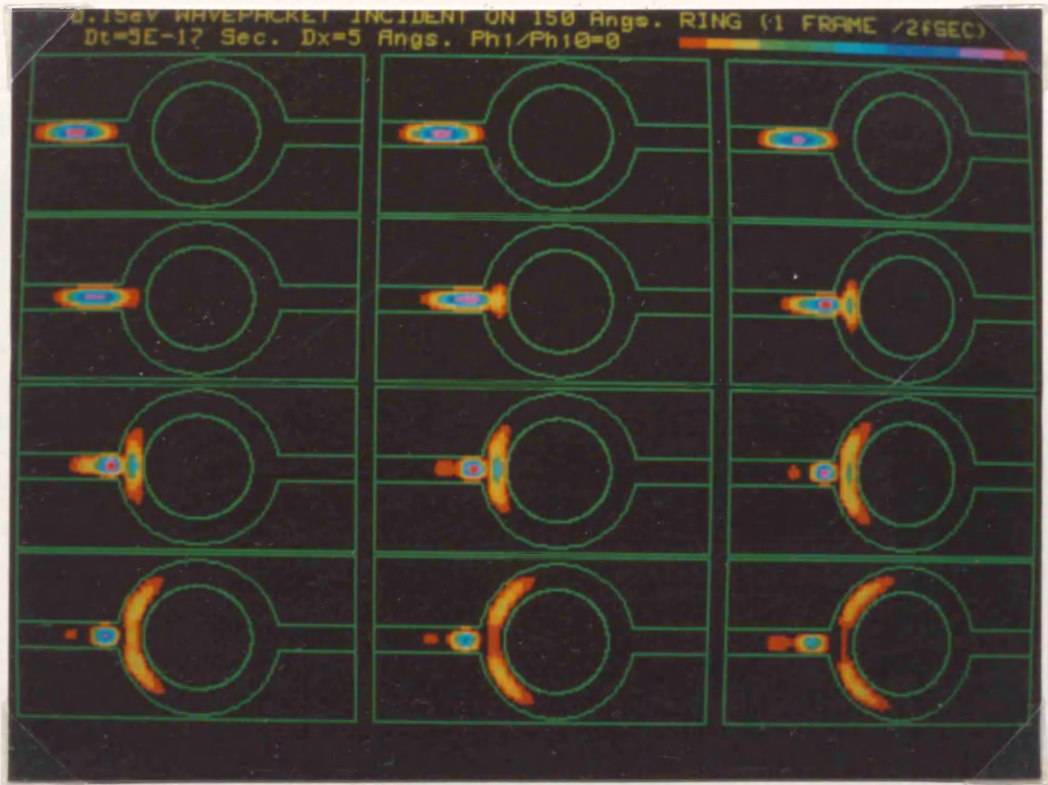


Fig 8.3

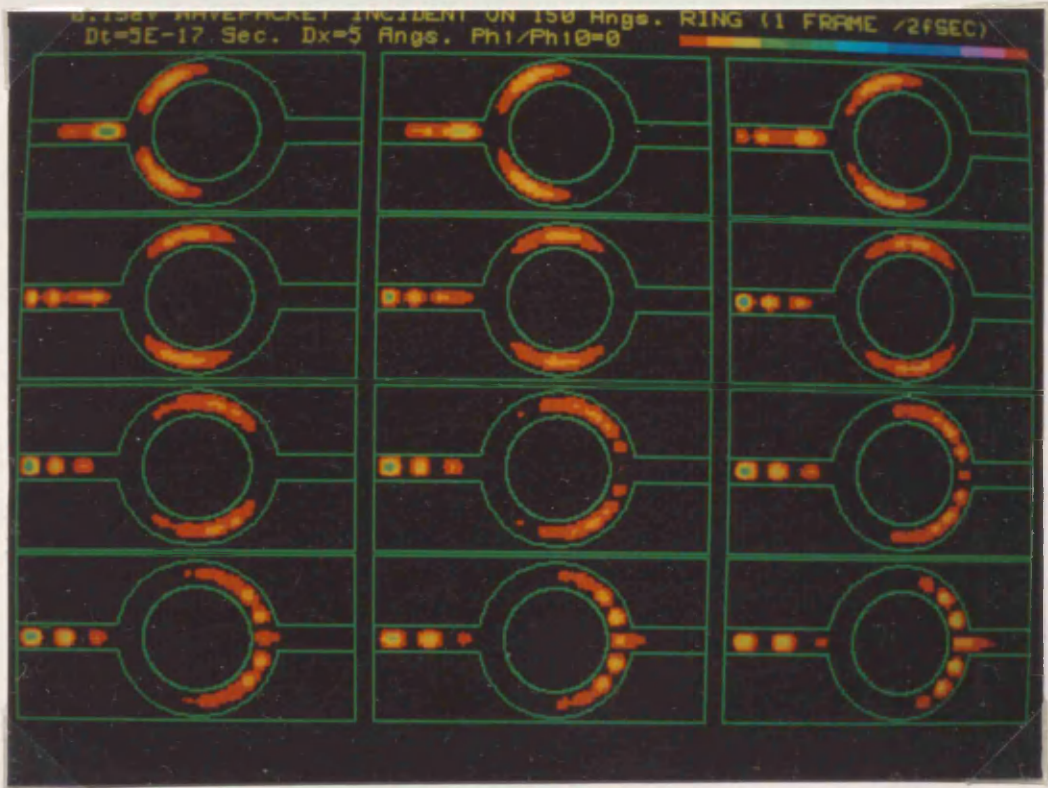
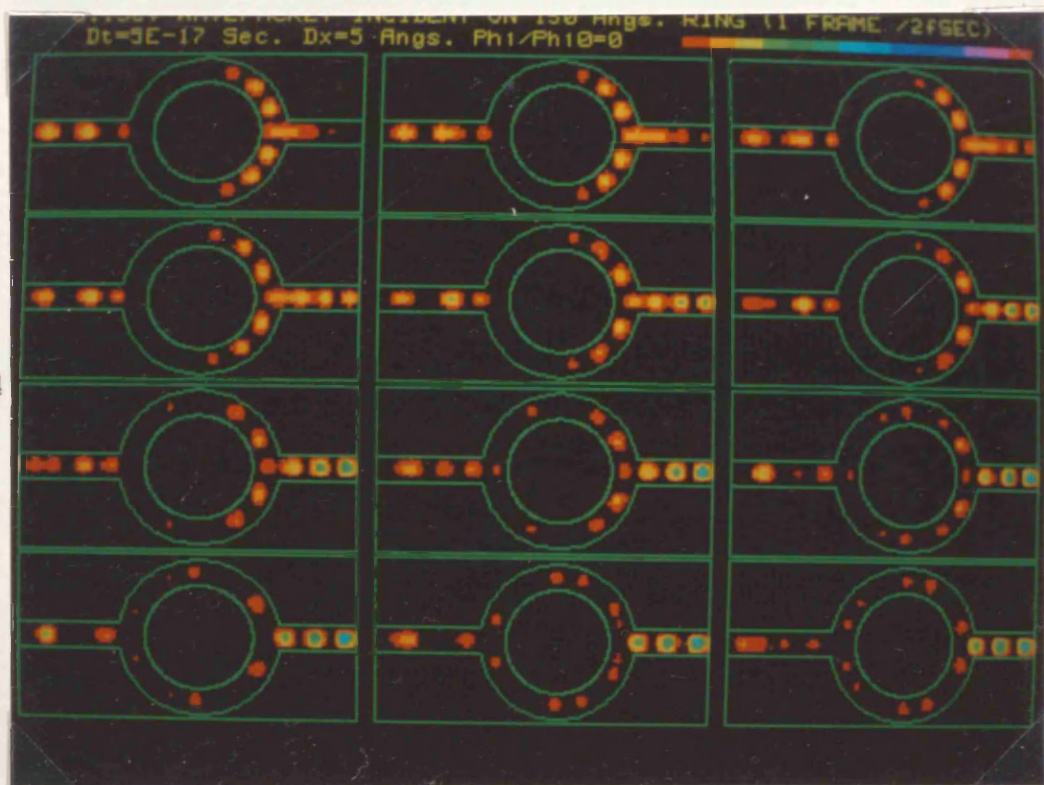


Fig 8.4



Figures 8.5 to 8.7: Propagation of a Gaussian wavepacket through a 30nm diameter test ring with a magnetic string of strength 0.5(h/e) threading the centre.

Fig 8.5

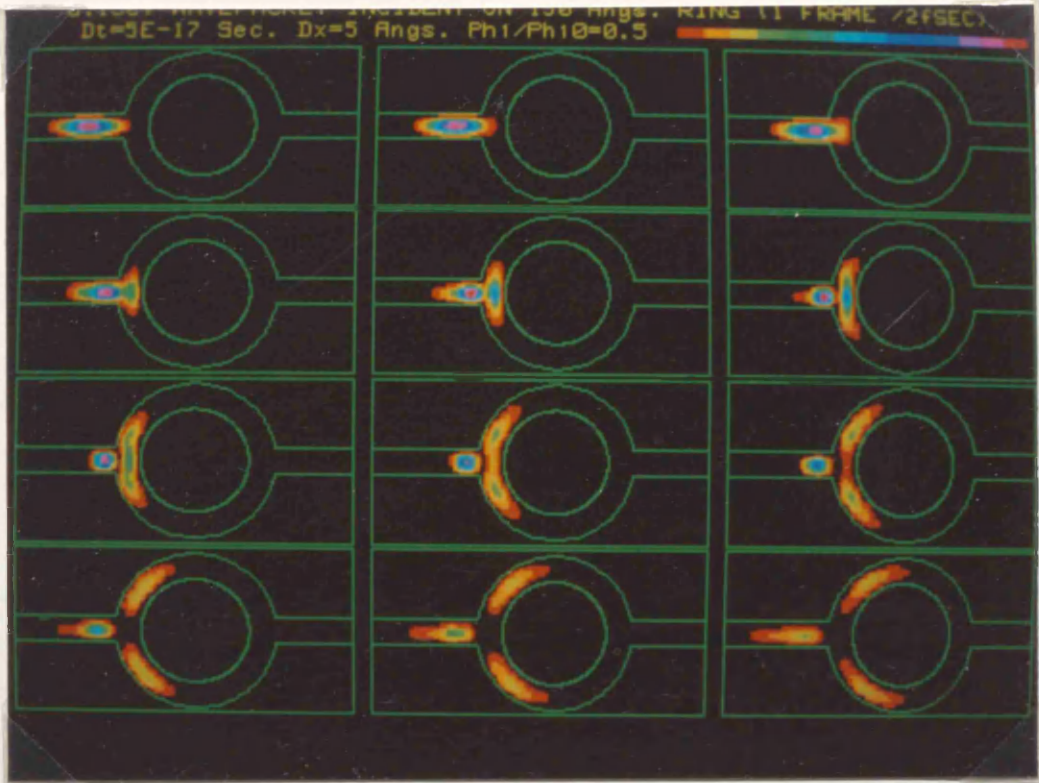


Fig 8.6

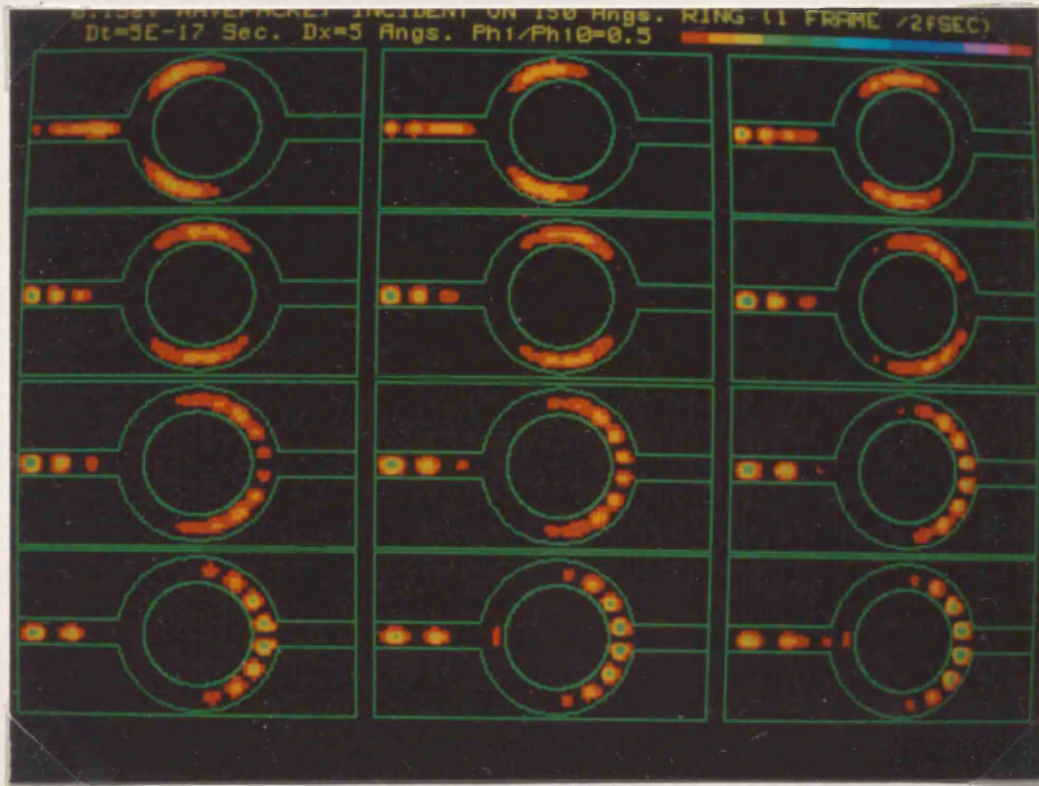
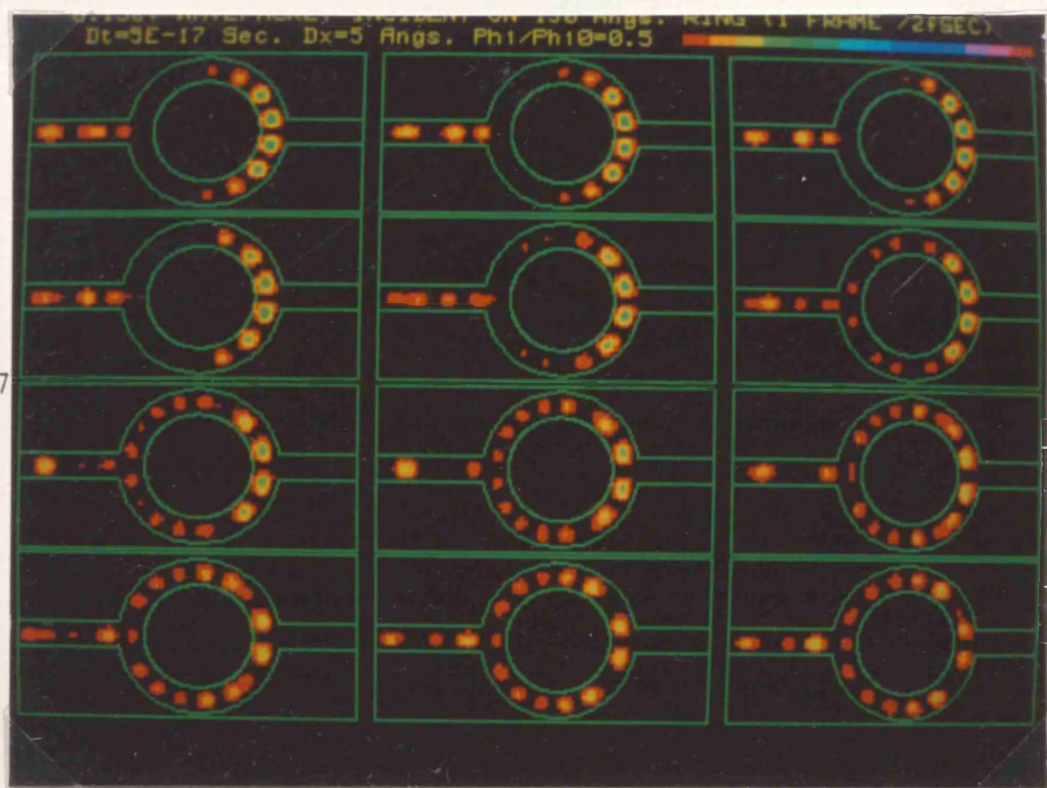


Fig 8.7



Figures 8.8 and 8.9 : 3D representations of the wavefunction in the absence and presence of a magnetic string (of strength $1/2(h/e)$) respectively.

Figure 8.8

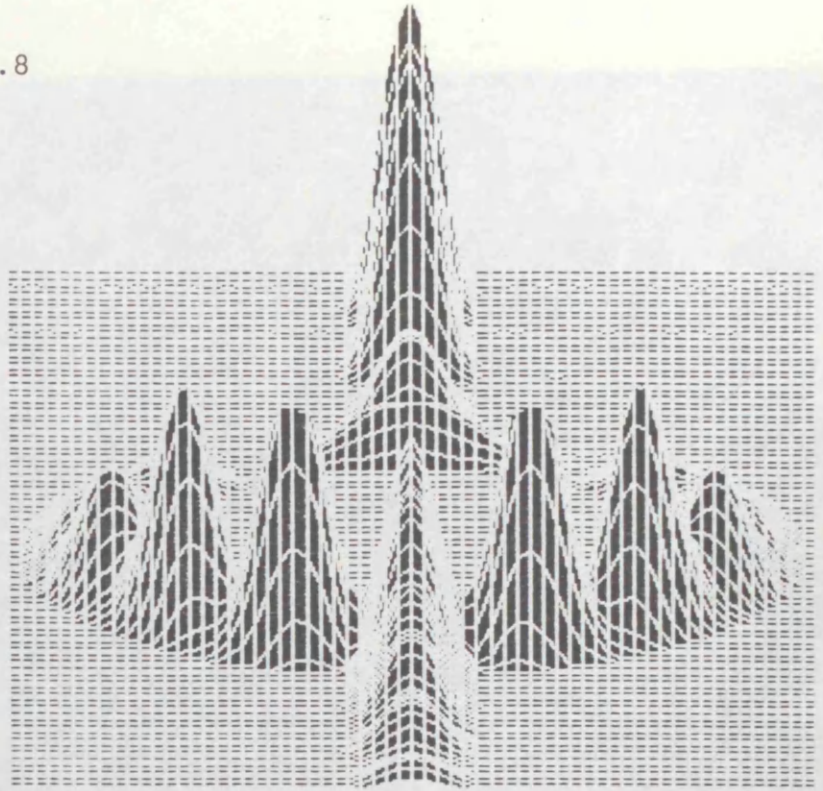
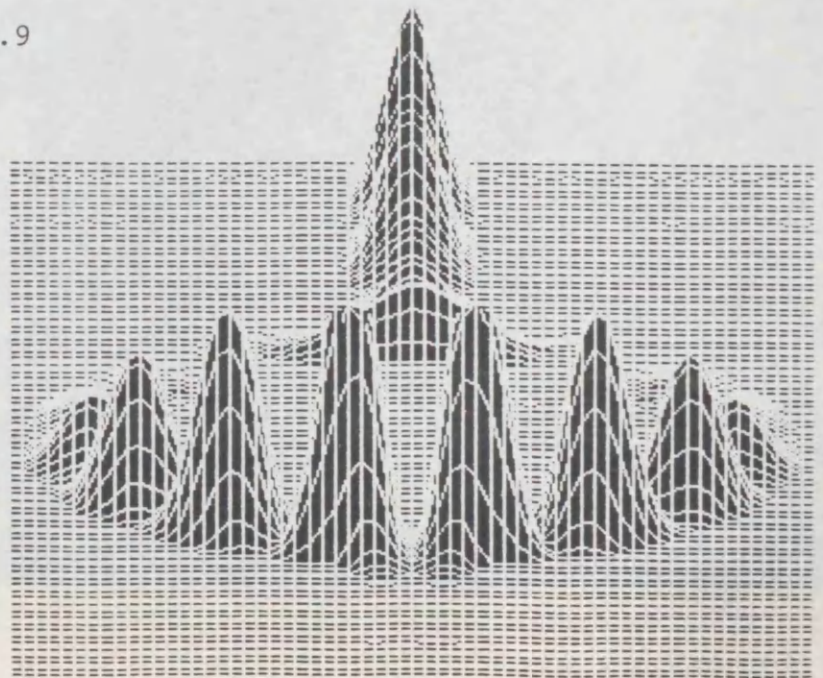


Figure 8.9



Figures 8.10 to 8.12: Propagation of a wavepacket through a 30nm diameter test ring in the presence of a uniform magnetic field with a flux of $0.5(h/e)$ threading the ring.

Fig 8.10

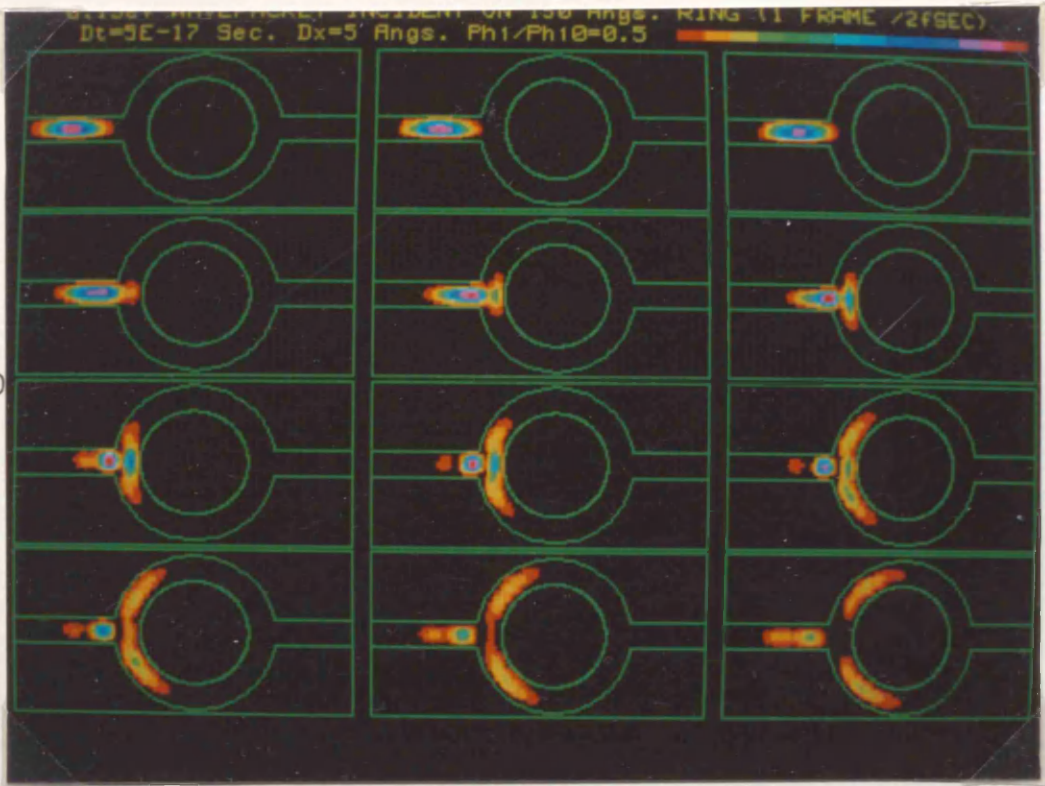
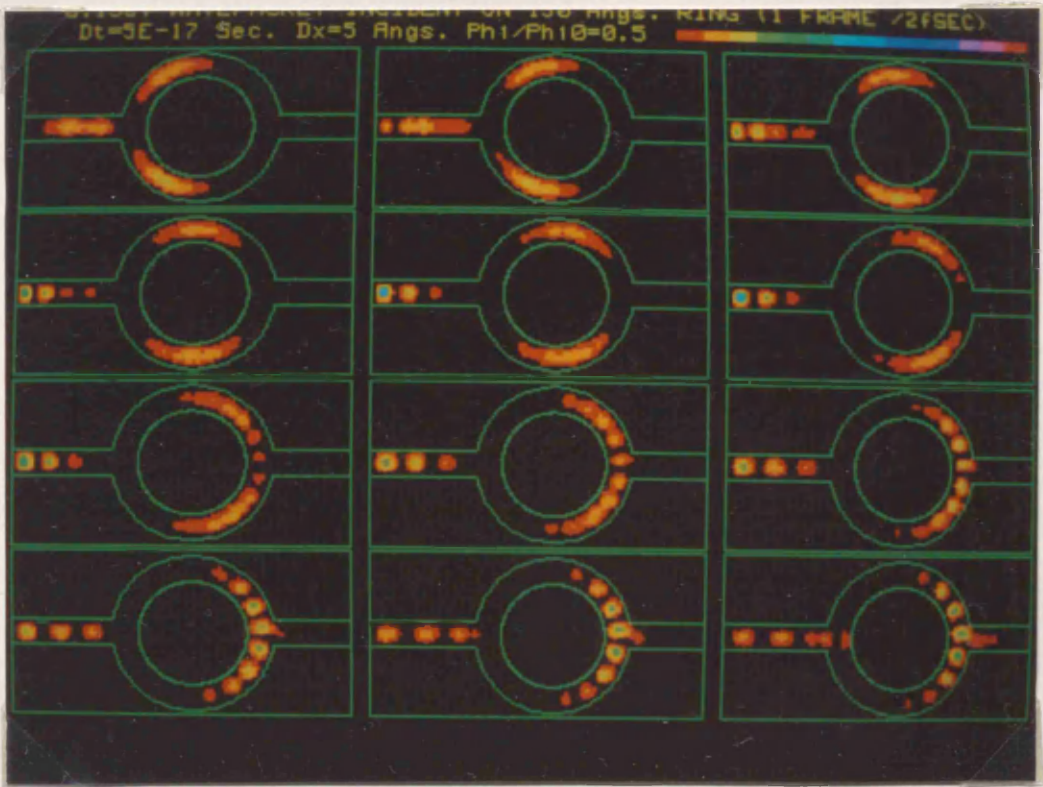


Fig 8.11



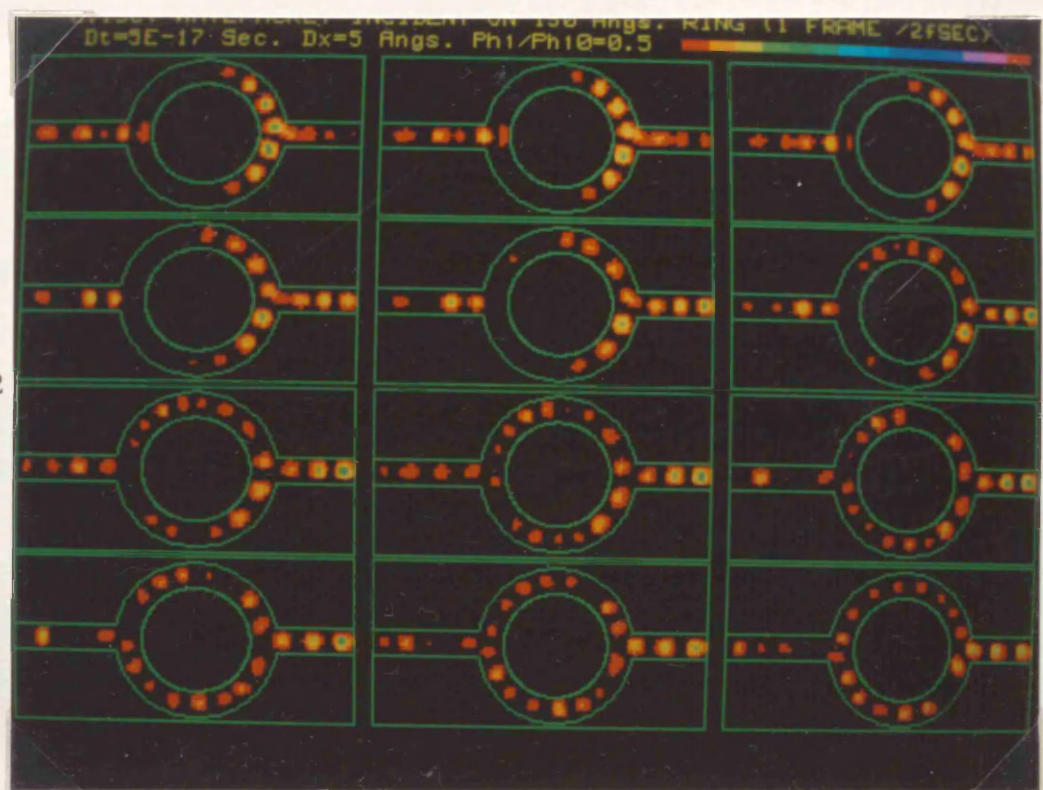


Fig 8.12

It is seen that although the position of the interference fringes change with respect to the overall envelope of the interference pattern, the envelope itself is shifted with respect to the output wire. This is caused by the Lorentzian force exerted on the electron. In this particular case the effect is exaggerated since the absolute value of the magnetic field required (to cause half a flux quantum to thread the ring) is relatively large compared to that required for rings of more realistic size. It does however, suggest that even in much larger practical devices, the amplitude of the conductance oscillations could become damped at high magnetic fields due to the overall envelope of the interference pattern progressively shifting until only secondary fringes are co-incident with the output. In the extreme case, the wavefunction could become totally localised in one arm of the ring with virtually no probability density co-incident with the output.

More sophisticated calculations on larger rings suggest that the transmitted part of the wavefunction could be expected to emerge in pulses due to multiple-transit effects around the ring (figures 3.13 to 8.33 for example). These pulses however, would exist on a much longer timescale than those seen due to reflections from the boundaries. One of the objectives therefore for the next series of simulations, was to model the boundaries or "contacts" more realistically.

8.4 A full-scale computer model.

The recent acquisition of a Floating Point Systems model 330 array processor by the theory group enabled a larger and hence more realistic simulation of an AB ring to be undertaken. The amount of memory available on the new machine allowed the number of points in both x and y-directions to be increased by a factor of four.

From typical sheet concentrations at heterojunction interfaces, it was estimated that the Fermi-energy was around 10meV (appendix C). It must be mentioned that in this work it is only the energy due to the longitudinal motion that is specified for the initial wavepacket and not the total energy. The total energy of the wavepacket in two dimensions (that which would be obtained by calculating the energy expectation value using the full 2D Hamiltonian) is this energy plus the energy due to the lateral confinement across the wire. For the wire widths considered in this project the lateral component of the energy is typically about 1meV.

Although there were no AB rings made by the shallow-etch technique available at the time this work was in progress, discussions with those in the department who were involved in fabricating the rings [J.McIntyre 1988], suggested that the minimum practical ring diameter that could be fabricated would be around 0.4-0.5 micron.

Taking into account the stability issues raised in chapter 7, a mesh size of 1.66nm was chosen, giving an outside diameter of 0.4 micron and a spatial discretisation error of around 0.5%. Accuracy and stability requirements using this mesh size indicated that a timestep interval of about 1fs would be adequate. The wire width was scaled up to 66.4nm wide, which was within the resolution limit of about 20nm for present day electron-beam lithography [Broers 1986].

To limit the amount of processor time and storage used, the simulation was run for a simulated time of 10ps, storing the values of the wavefunction at each mesh point every 0.5ps. A simulation for a longer time was performed, but no significant features became apparent in addition to what had already been observed in the first 10ps.

The array processor used for these numerically intensive computations did not have direct access to a graphics screen, which meant that data from the simulation had to be written to disk and then copied to another machine that did have a graphics terminal. A disadvantage of the system in use at the time was that the data had to be copied using a standard transmission link, and it was found that the time taken to copy data for the real and imaginary parts of the wavefunction (in double-precision arithmetic) was unacceptably long (about 10-15 minutes per frame). Since it was the probability density which was of primary interest, the amount of stored data was reduced in the first instance by only storing the square-modulus of the wavefunction, straightway saving a factor of two on storage. When the data were plotted as a colour contour map, each numeric value at a particular mesh point was assigned an integer between 0 and 60 to represent the colour of that mesh cell. Therefore, in order to further reduce the storage requirement, the values of the probability density were pre-coded into the appropriate integers before storage. A final step that was taken was to use a simple data compression technique to encode large areas of constant colour values into a single integer (larger than 60). The use of all these techniques resulted in a copy-out time of a few seconds and an appreciable saving (about 10-fold) in disk space, although this was not originally of primary concern.

Once the data had been read from disk onto the VAX computer, the data decode time was in the order of about 30 seconds, after which the decoded data were plotted on a high-resolution graphics screen using GKS-standard graphics software.

In the large-ring simulations only the case of a uniform magnetic field distribution was considered because in using a larger number of mesh points a small instability was observed in the simulations involving the magnetic string distribution. The instability was apparent in the gradual increase in the wavefunction normalisation in the closed system (without absorbing contacts). Whilst the error was not very large for each timestep, the large number of timesteps required (about 10,000) to determine the transmission and reflection coefficients meant that the error in the norm accumulated to about 5-10% over the entire simulation. The reason for this is most likely due to numerical round-off errors arising from the polar to Cartesian co-ordinate conversion which must be performed when setting-up the appropriate vector potential, since the instability was only observed when the larger grid was used.

Whilst the simulation of the propagation of the wavepacket could have been performed by renormalising the wavefunction at each timestep, an accurate determination of the transmission coefficient by measurement of transmitted currents into the contacts would not have been possible because one would not know how much of the change in the norm was due to absorption into the contacts and how much was due to the error. In view of this, and the fact that only a finite amount of computer resources were available, it was decided to concentrate on one of the field distributions (the uniform field case) for both the contour plots of the probability density and the magneto-transmission results. Thus all results would be consistent with regard to the magnetic field distribution. In any case the uniform field case is nearer to experimental reality than the magnetic string distribution, so the uniform-field results would be more appropriate for comparison with experiment.

8.4.1 General analysis of results for a large ring.

Figures 8.13 to 8.33 show a time sequence of frames from the simulation of the larger ring described above taken at intervals of 0.5ps. The position probability density is again plotted on a linear scale consisting this time of 60 colours. In these simulations, absorbing contacts modelled by regions of complex potential were incorporated as discussed in section 7.8 and are delineated by the vertical lines at the edges of the simulated region.

The effect of the contacts was to greatly reduce the reflection at the boundaries, although some reflection was still apparent, indicating that the wavefunction did not propagate as it would have done if completely free. However, it has been remarked [Thouless 1980] that in reality one cannot expect a wavepacket to continue broadening indefinitely, and some re-localisation of the wavepacket must occur at some point due to inelastic collision processes.

The most striking difference between these and earlier results was that the form of the wavefunction was much more complex. The characteristic scale of the fragments of probability density was found to be about 24nm, which is about half the DeBroglie wavelength of a 10meV plane wave.

The highly intricate nature of the wavefunction, is suggestive of quantum chaotic behaviour similar to that found by McDonald and Kaufman [1979] (figure 8.35) in their solution of the time-independent Schrodinger equation for a stadium-shaped region which is known to be classically chaotic. Figure 8.34 shows a plot of the probability density for the ring described above with the imaginary part of the potential set to zero, so that the boundaries were totally reflecting. This simulation was run for simulated time of 50ps which gave ample opportunity for the initial wavepacket to interact with all the boundaries and to distribute itself throughout the ring. A calculation of the energy eigenvalue spectrum would be of interest here, but the difficulty in diagonalising such a large Hamiltonian matrix precluded it from being undertaken at this time. In the work of McDonald and Kaufman, it was found that the non-chaotic circular system had an eigenvalue distribution (as a function of level separation) which was roughly exponential in shape. Small separations between levels were preferred to large ones, giving rise to level clustering. In the stadium problem, the distribution was fundamentally different, with the most favoured separation being some intermediate energy, giving rise to a level-repulsion effect.

Figures 8.13 to 8.33: Propagation of a Gaussian wavepacket through a 400nm diameter ring with no applied field.

Fig 8.13

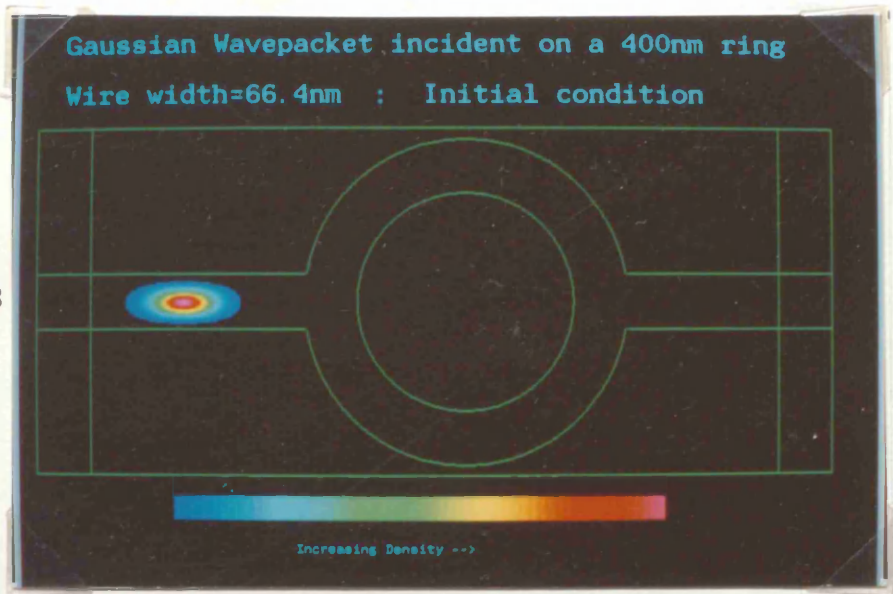


Fig 8.14

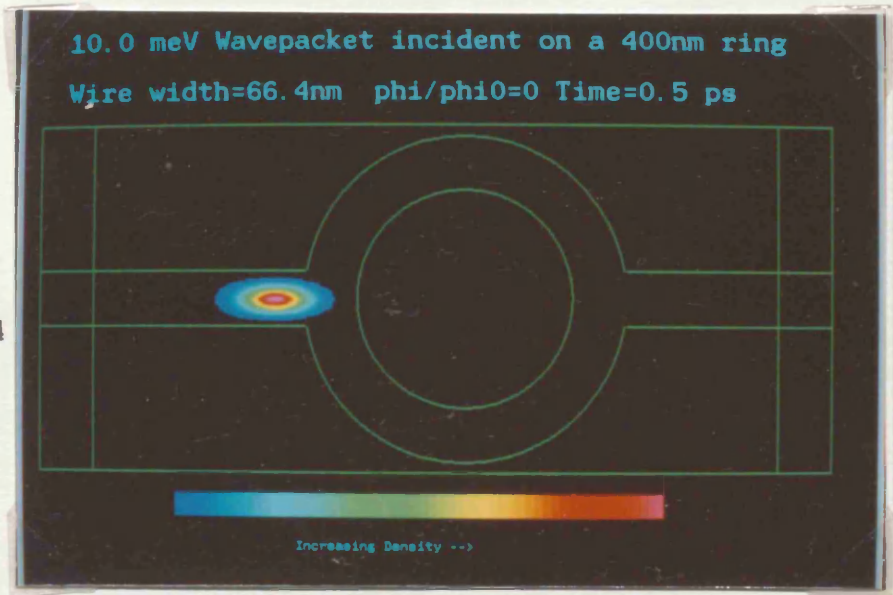


Fig 8.15

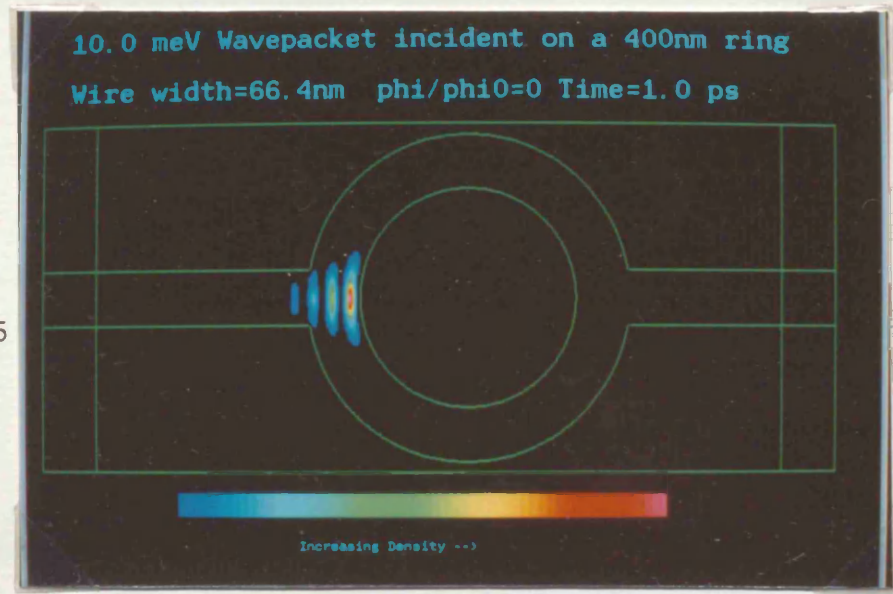


Fig 8.16

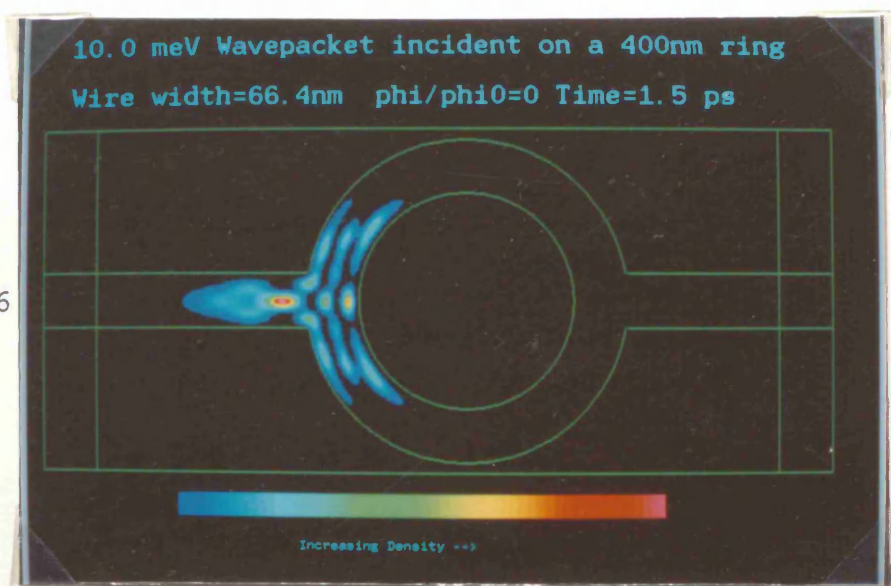


Fig 8.17

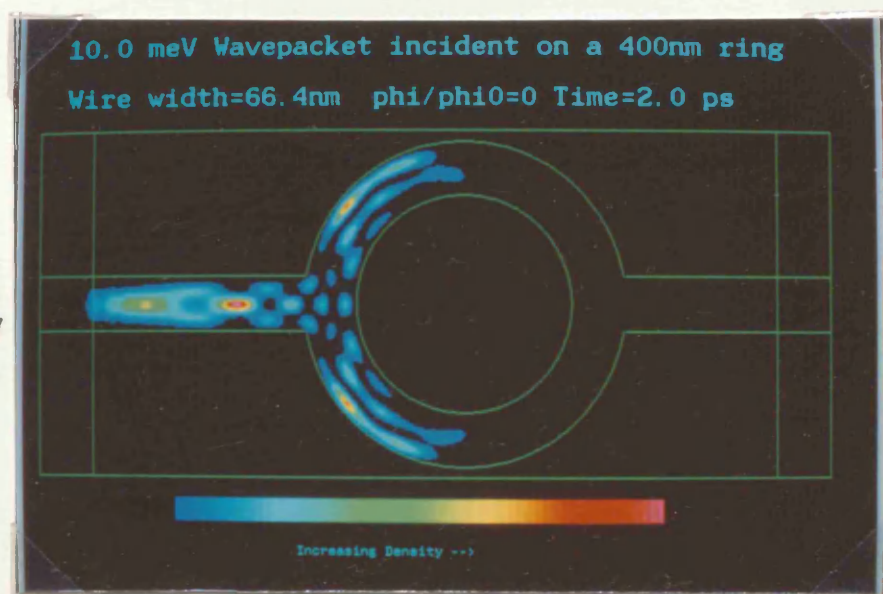


Fig 8.18

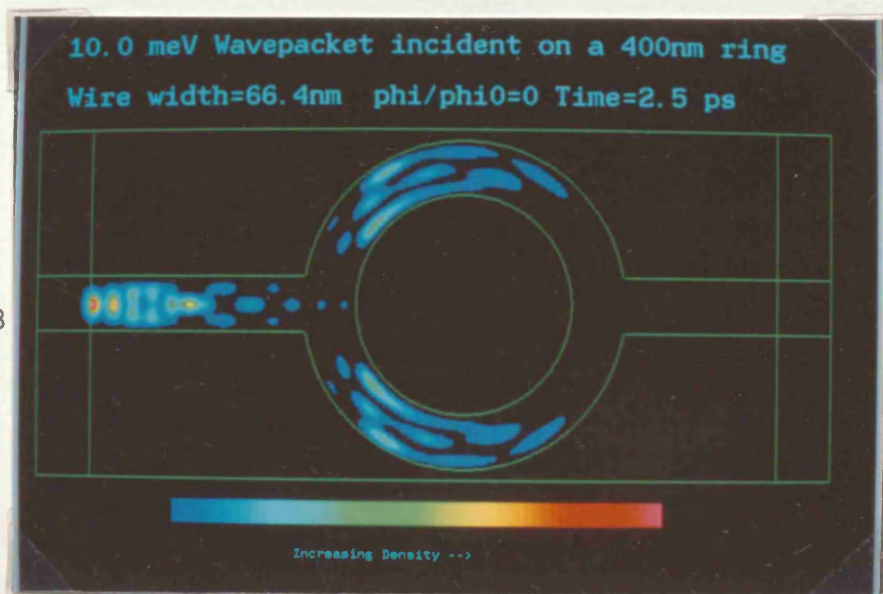


Fig 8.19

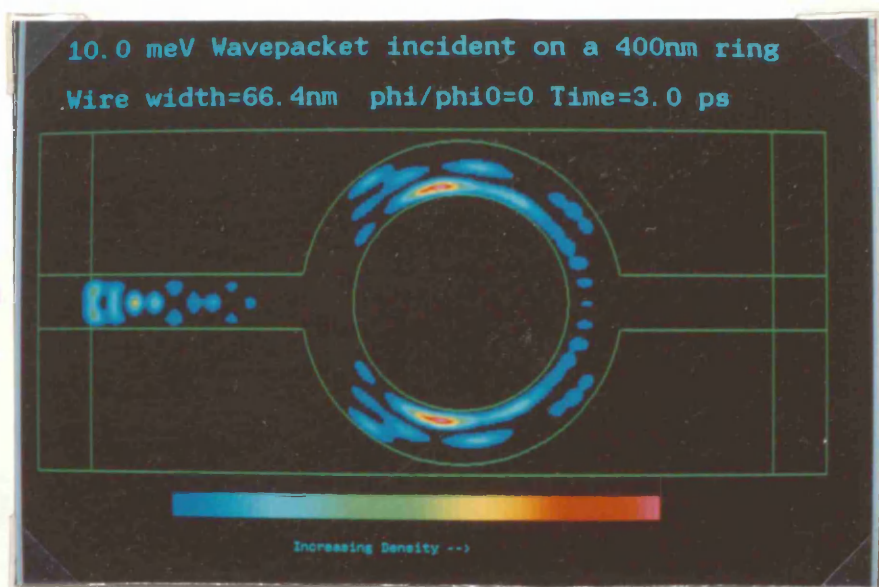


Fig 8.20

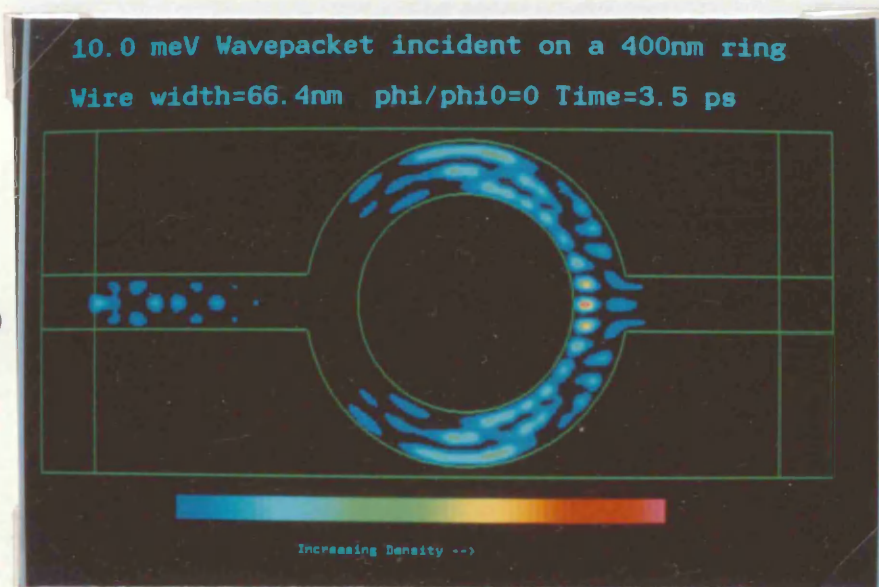


Fig 8.21

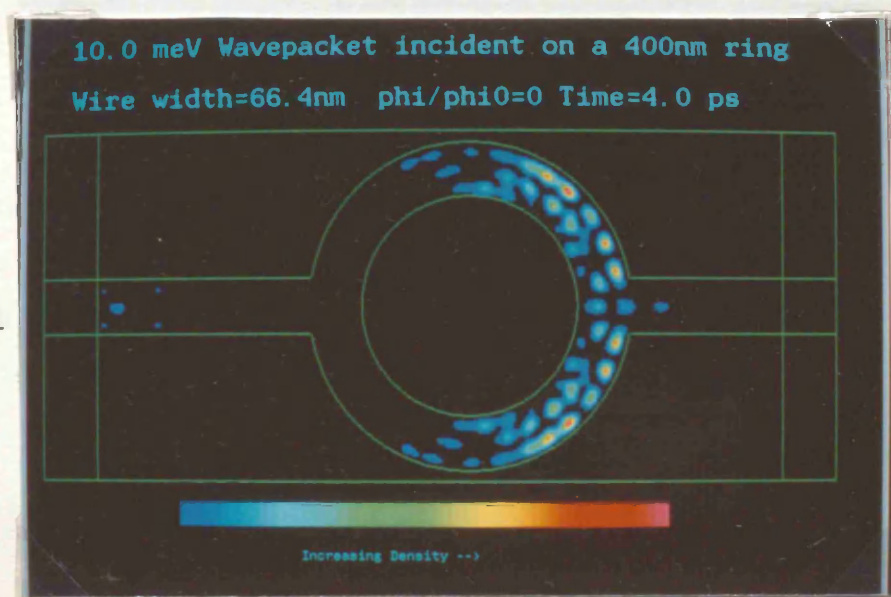


Fig 8.22

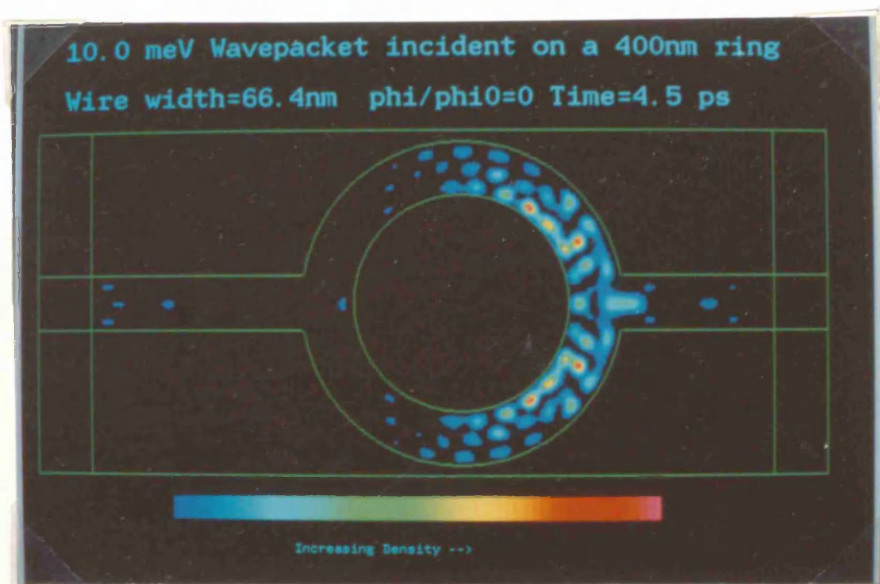


Fig 8.23

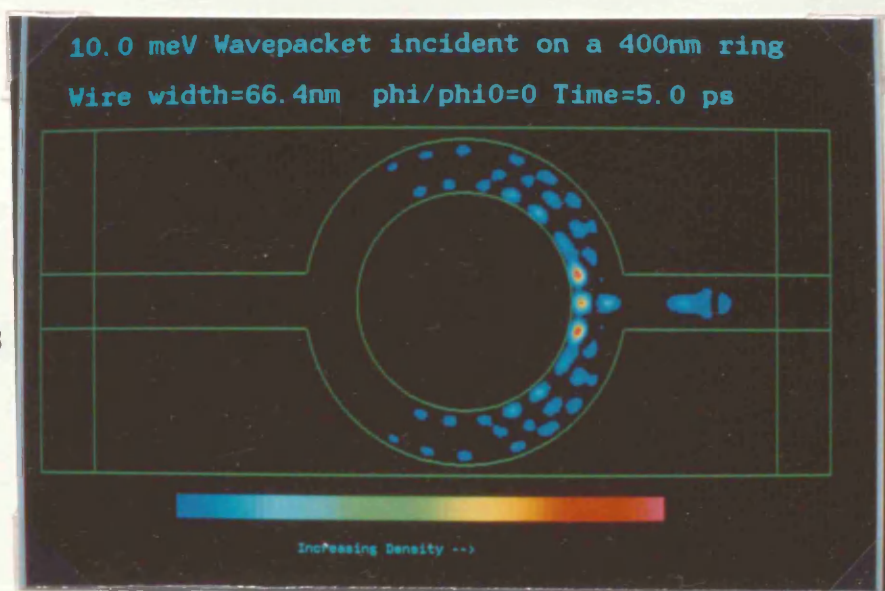


Fig 8.24

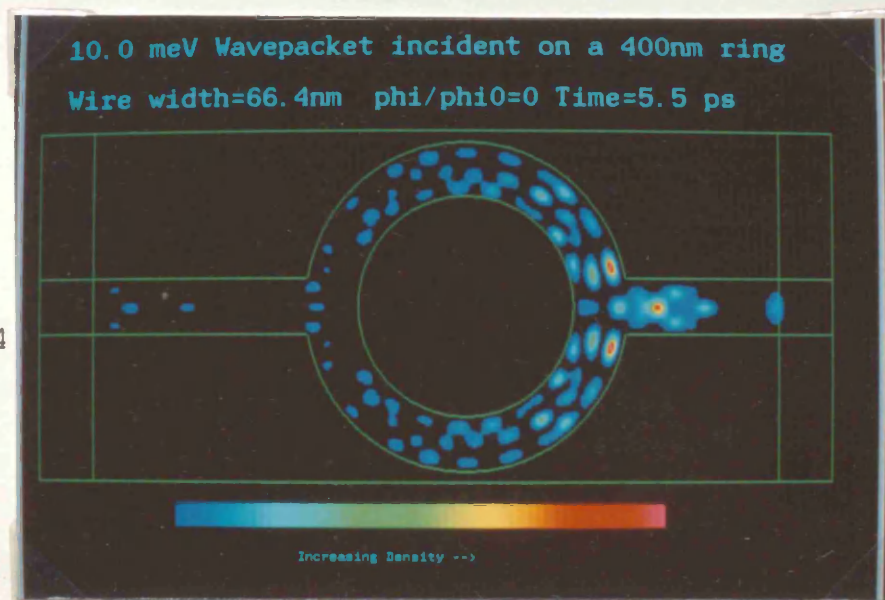


Fig 8.25

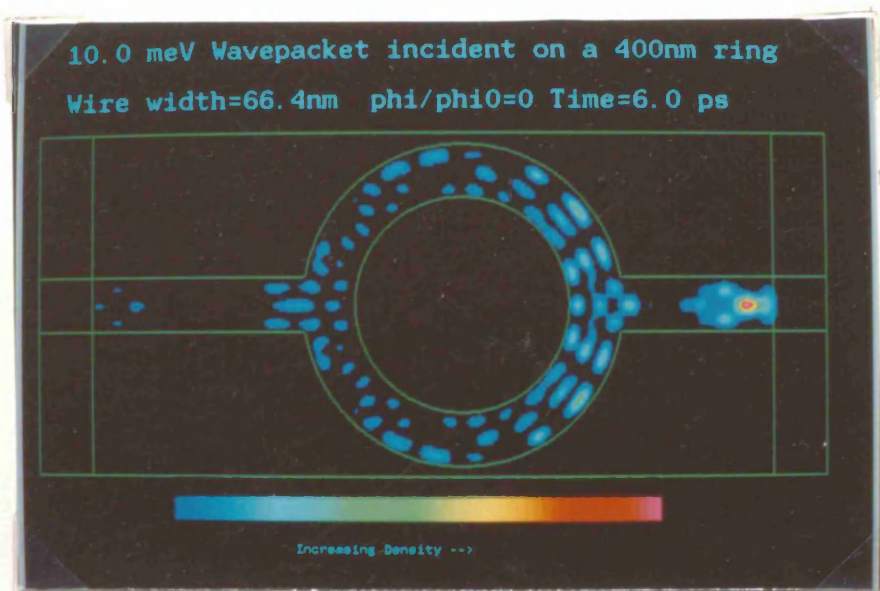


Fig 8.26

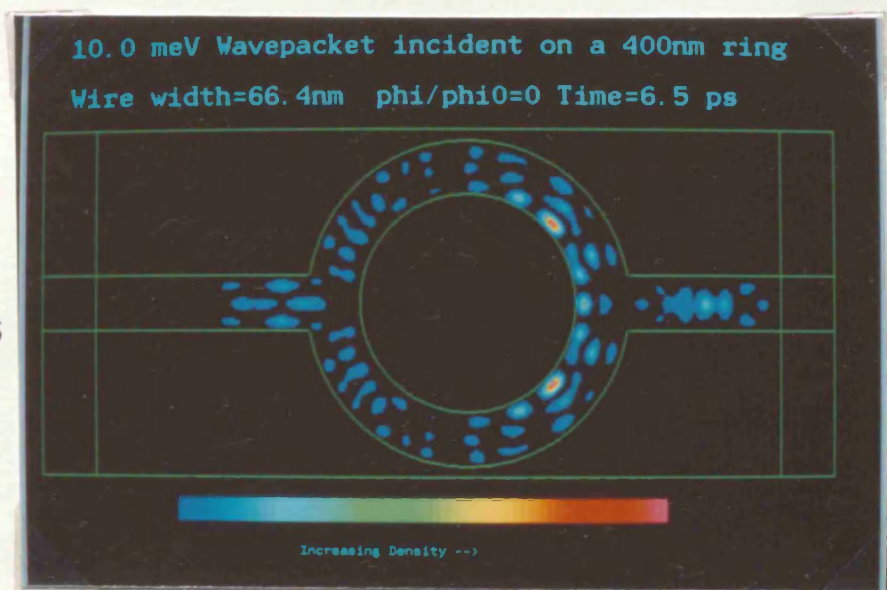


Fig 8.27

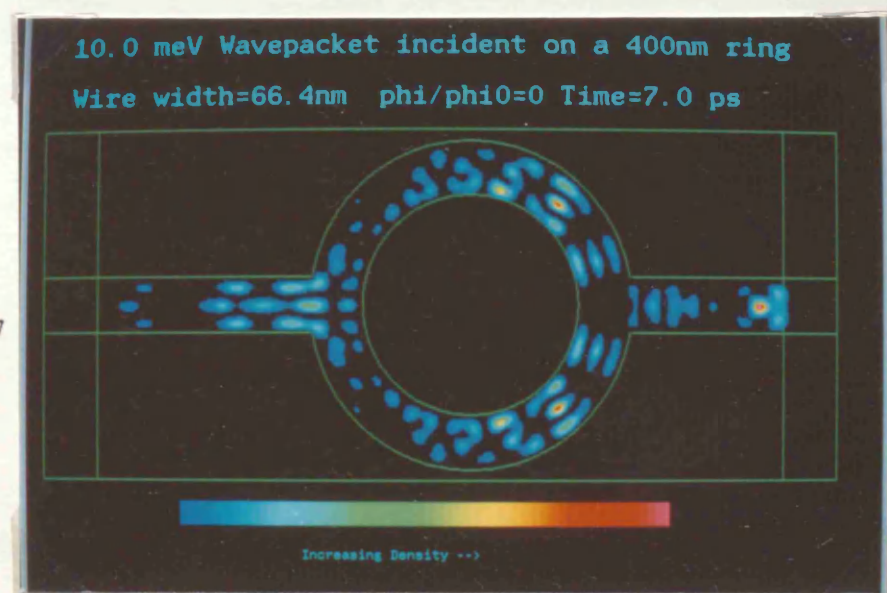


Fig 8.28

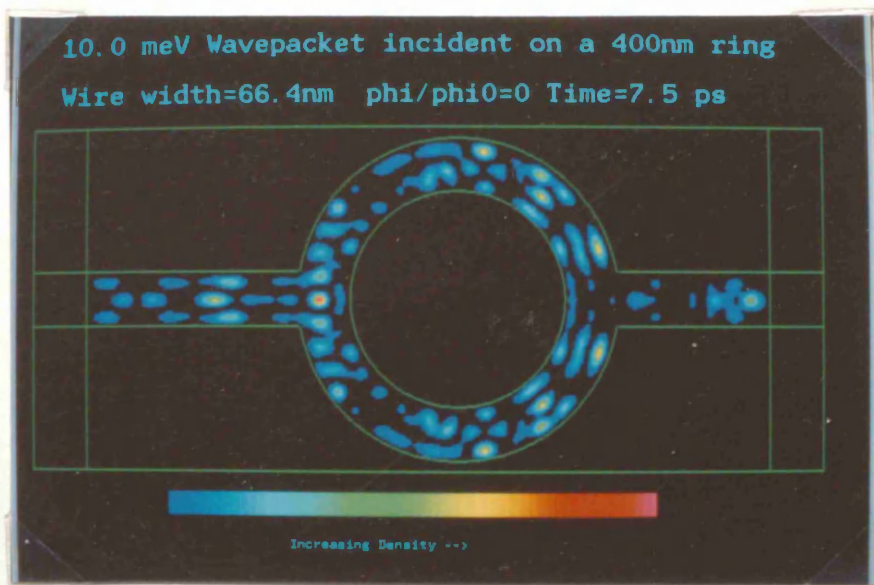


Fig 8.29

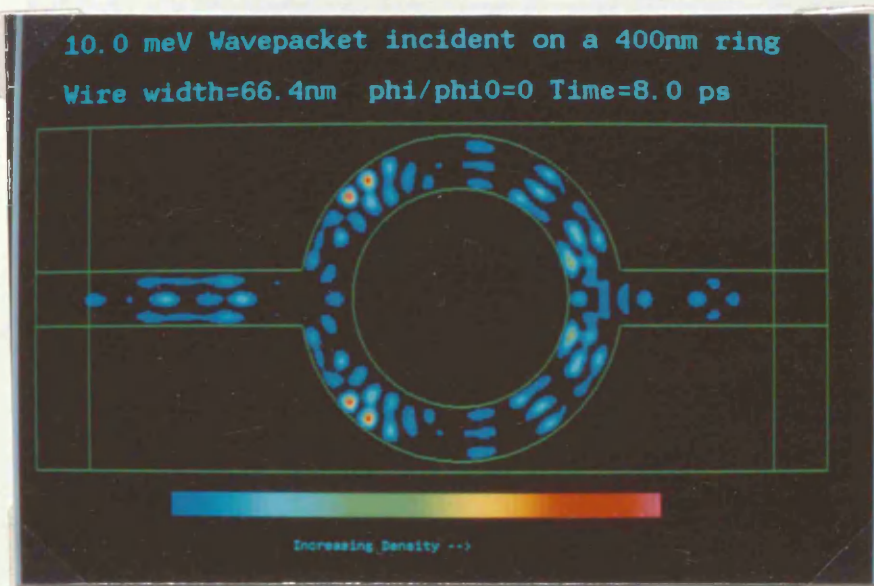


Fig 8.30

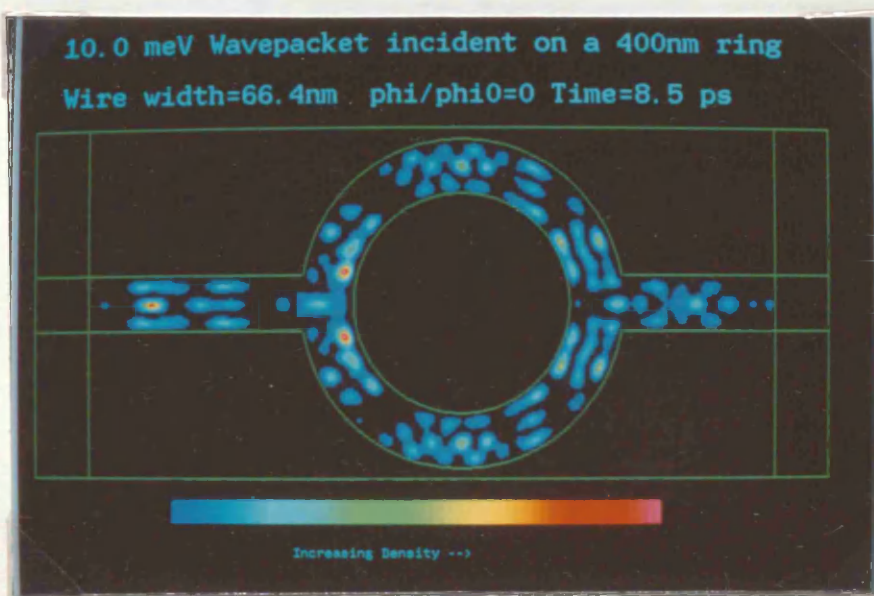


Fig 8.31

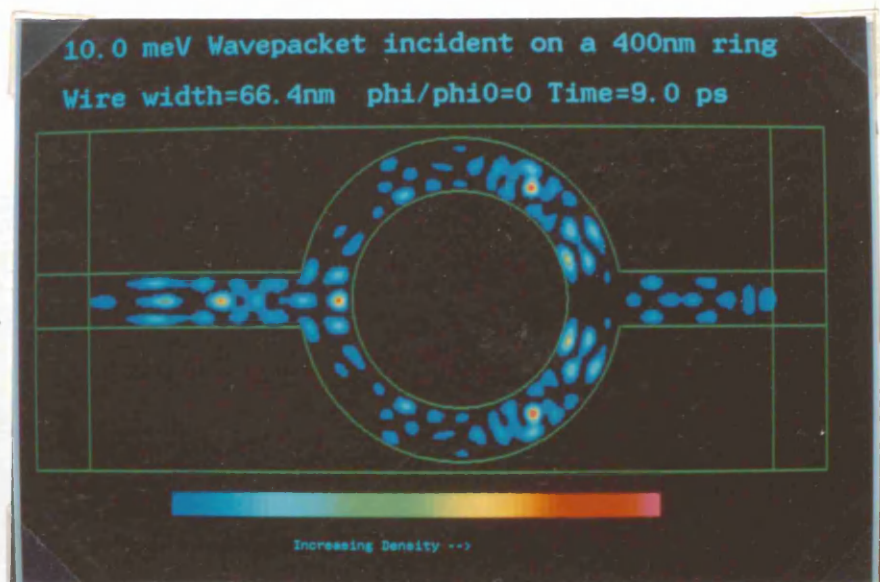


Fig 8.32

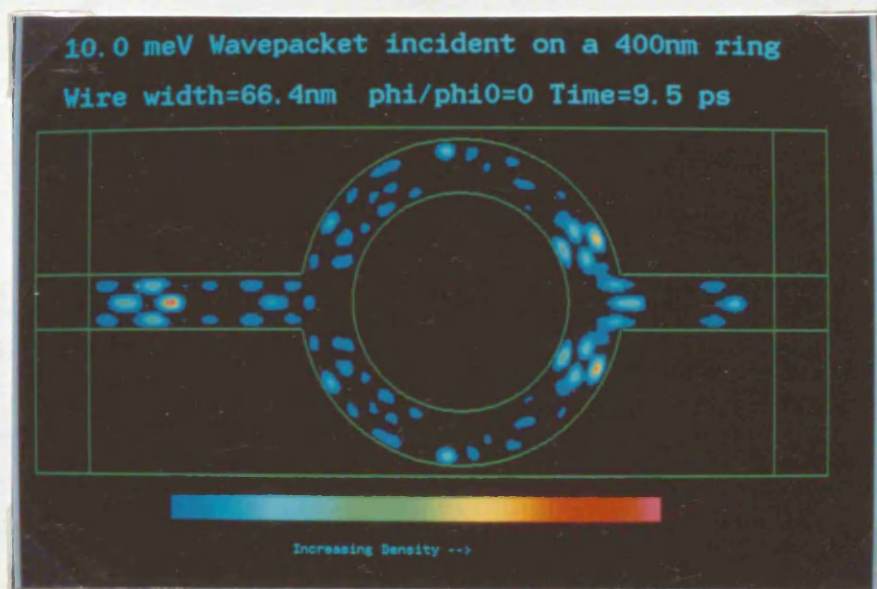


Fig 8.33

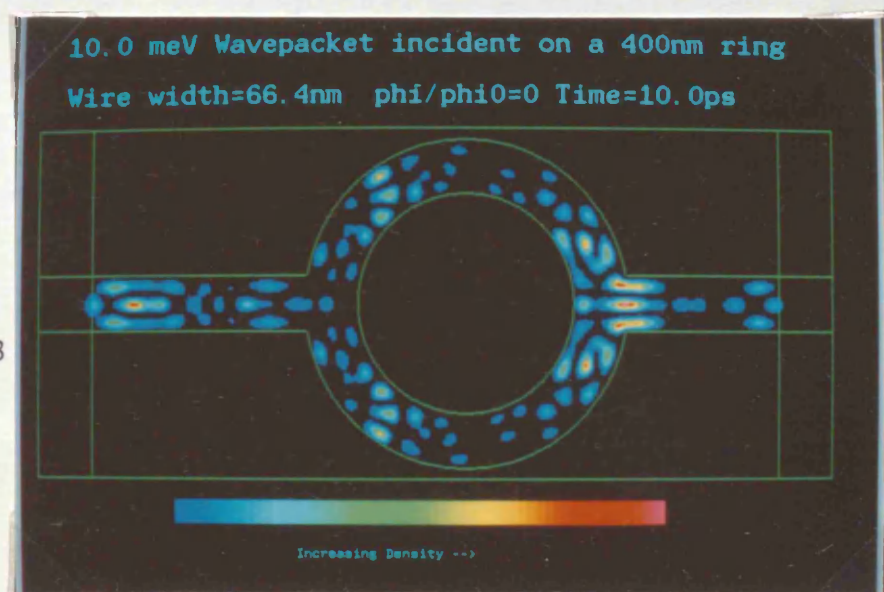


Figure 8.34 Complex form of the wavefunction when the ring is modelled as a closed system. The probability density in this plot is coded in the sequence: red-blue-green-yellow-orange in order of ascending magnitude.

10meV Wavepacket incident on a 4000A ring
dt=1. 0E-15 sec dx=16. 0E-10 m phi/phi0=0

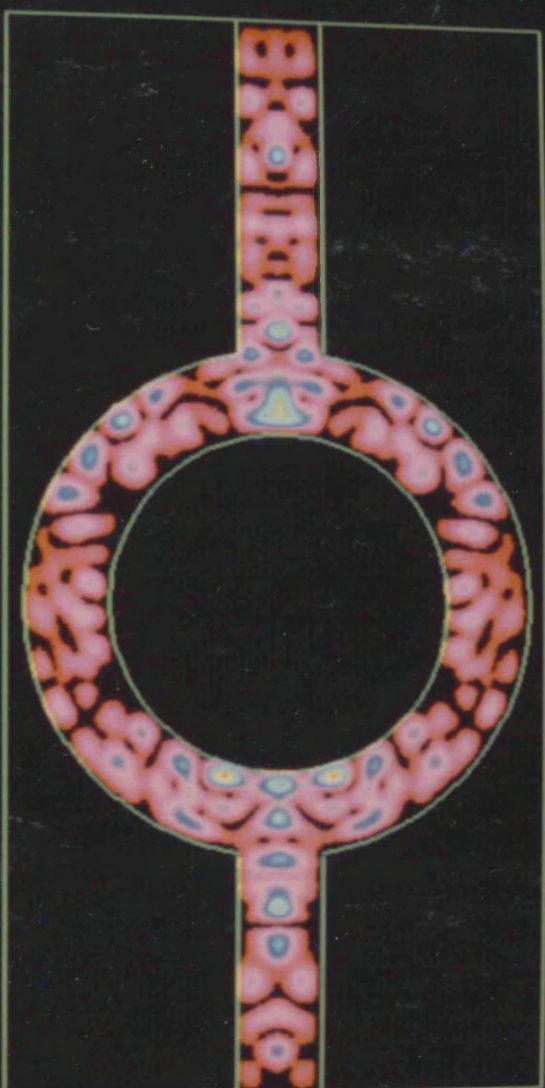
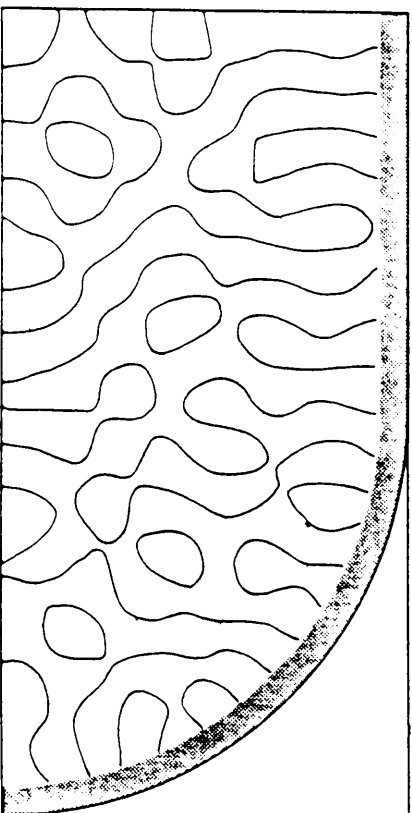


Figure 8.35 Nodal contours of $\psi(x,y)=0$ for one quadrant of a stadium-shaped region arising from the solution of the time-independent problem. The orientation of the nodal contours appear to random.



Figures 8.36 to 8.55: Propagation of a Gaussian wavepacket through a 400nm diameter ring with a uniform field applied.

Fig 8.36

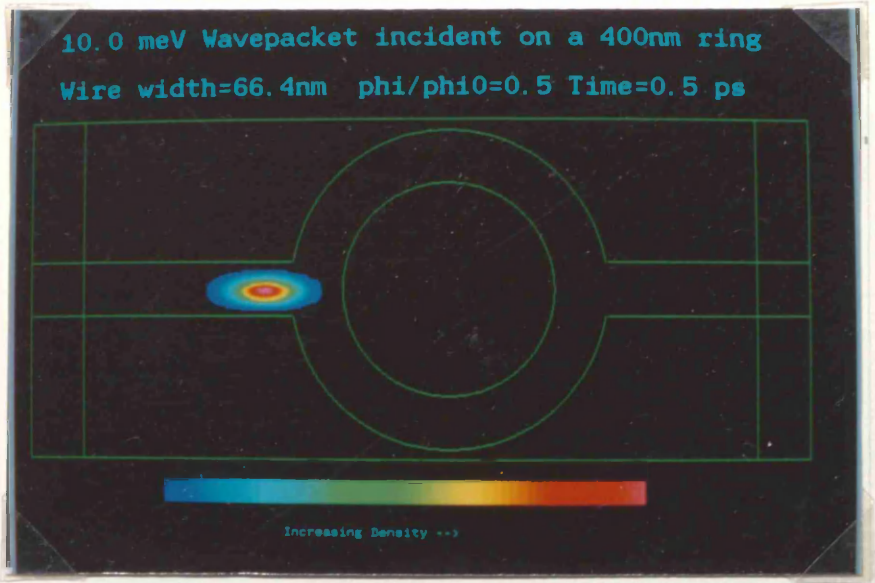


Fig 8.37

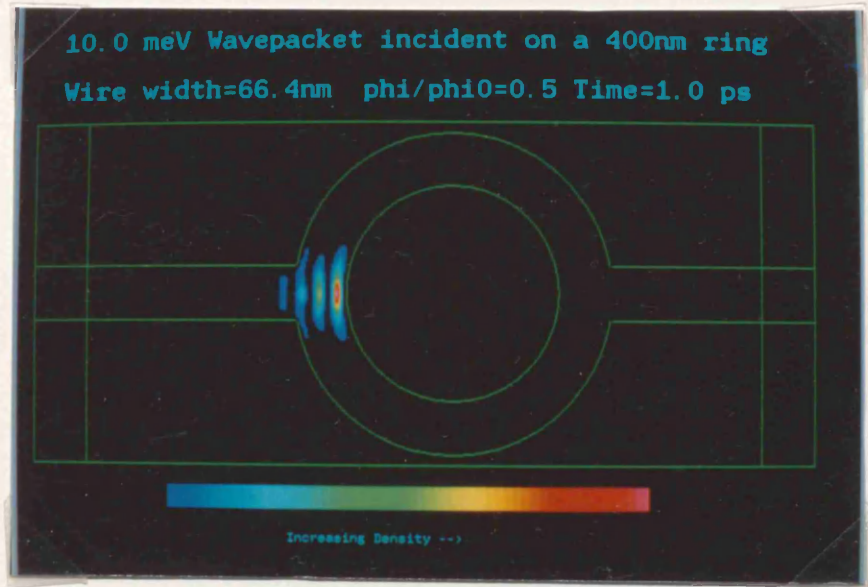


Fig 8.38

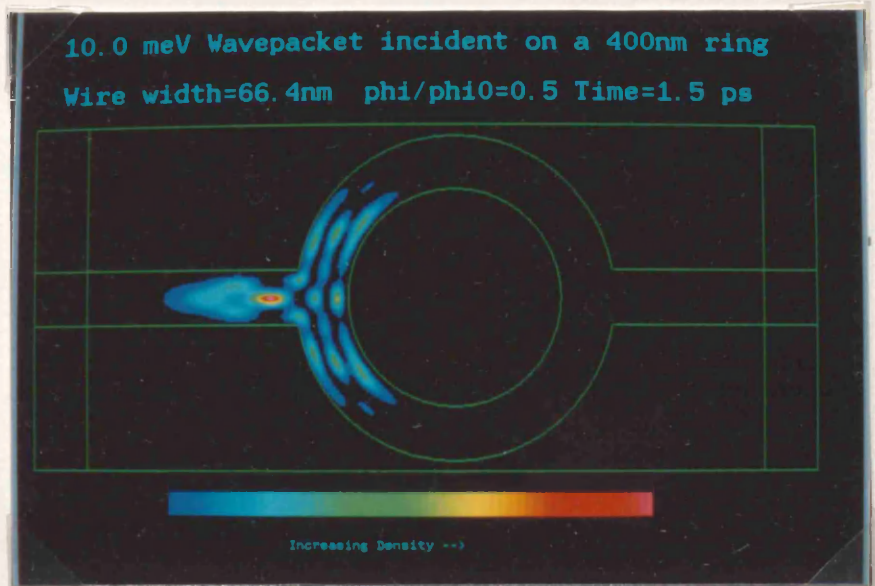


Fig 8.39

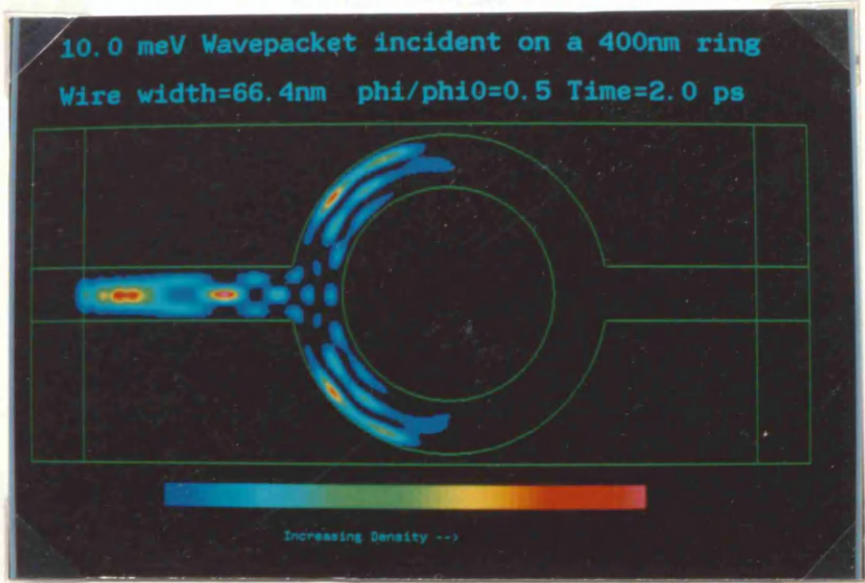


Fig 8.40

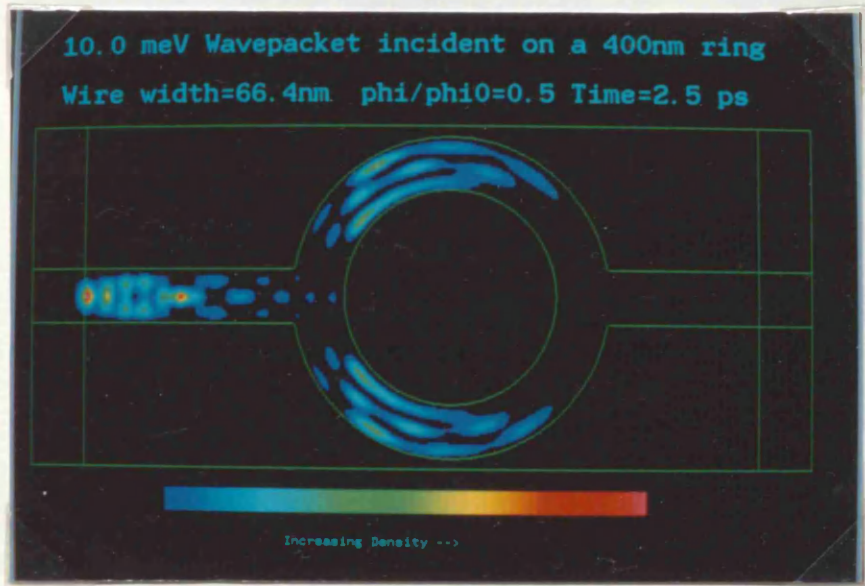


Fig 8.41

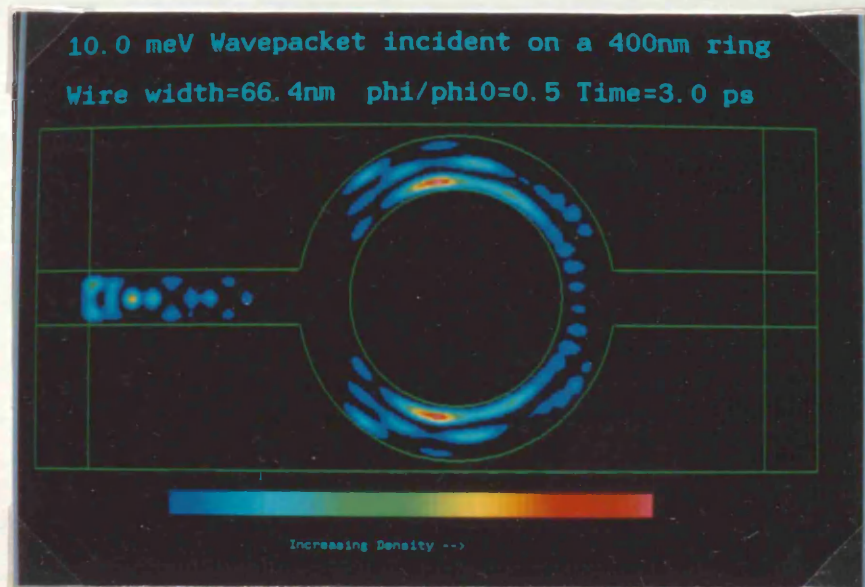


Fig 8.42

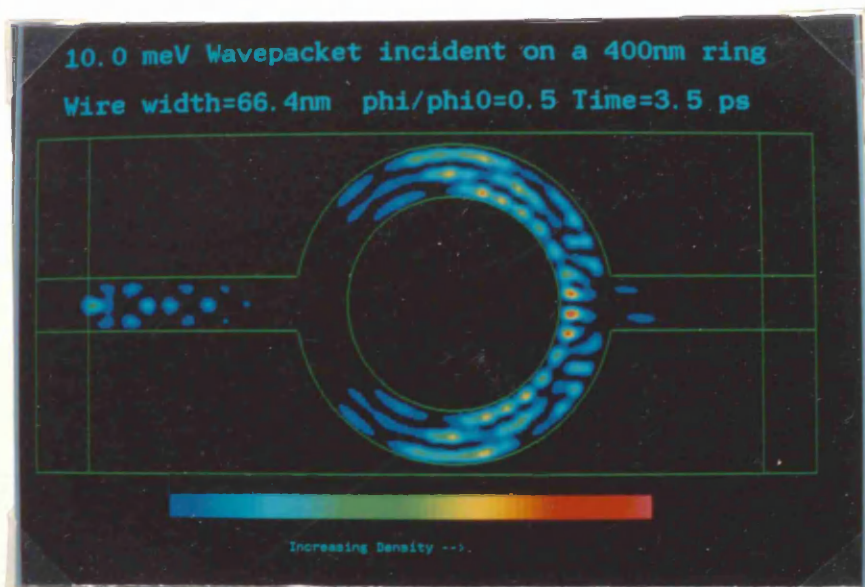


Fig 8.43

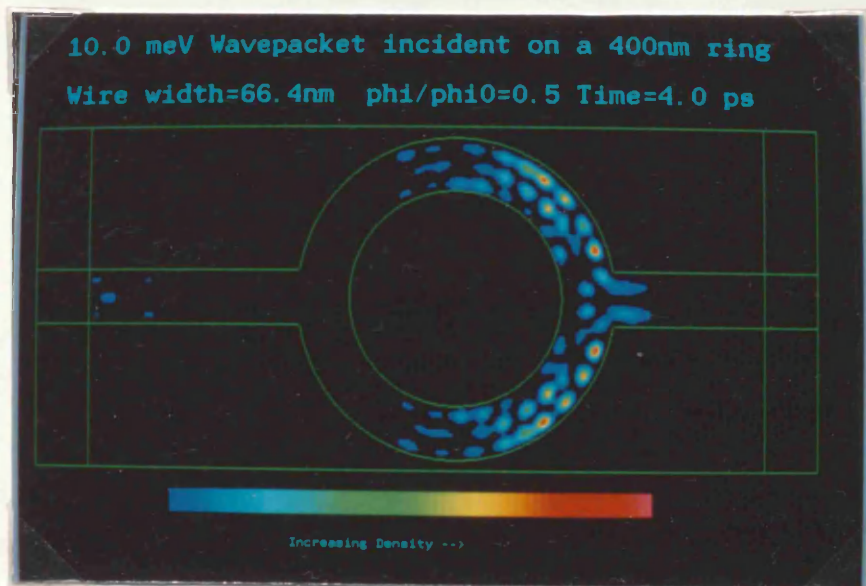


Fig 8.44

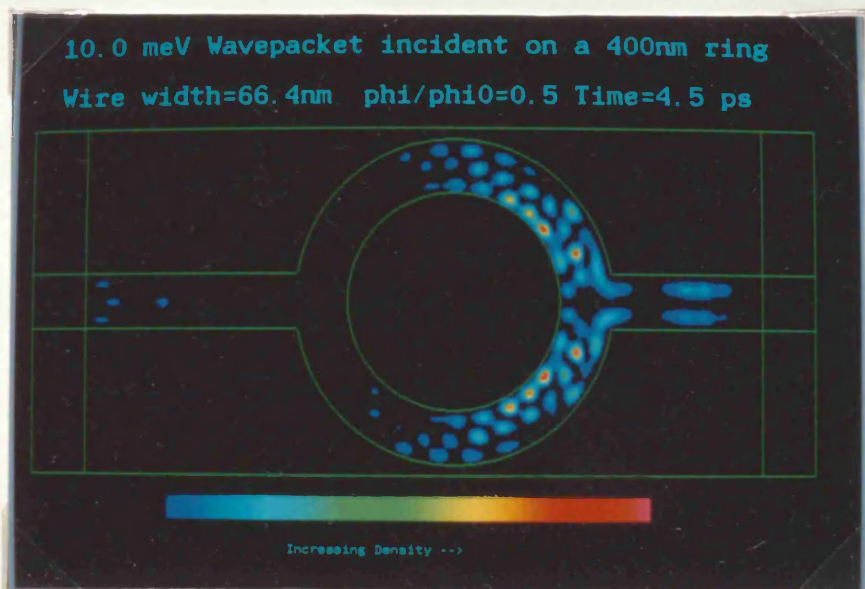


Fig 8.45

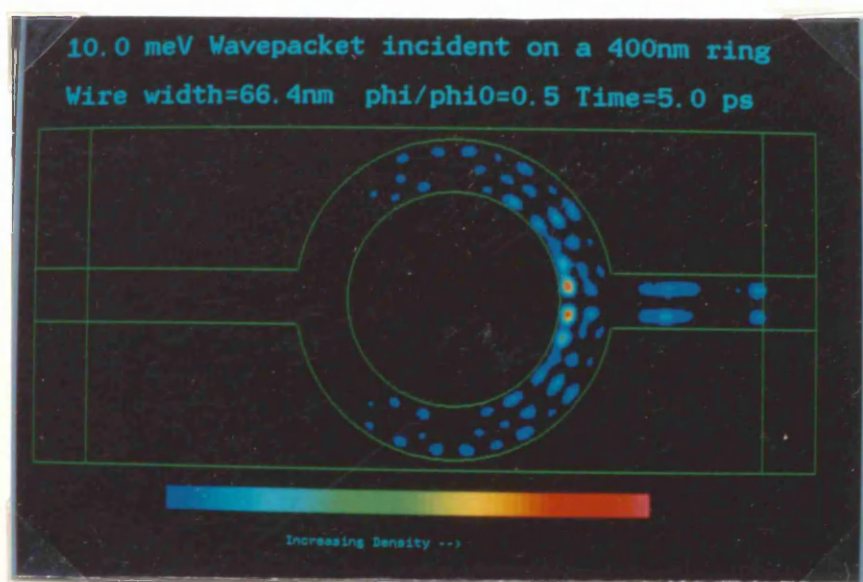


Fig 8.46

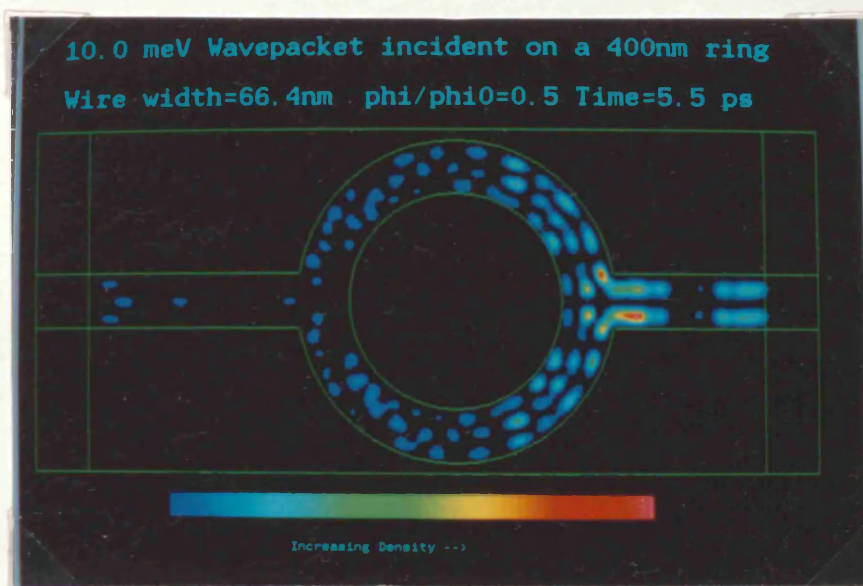


Fig 8.47

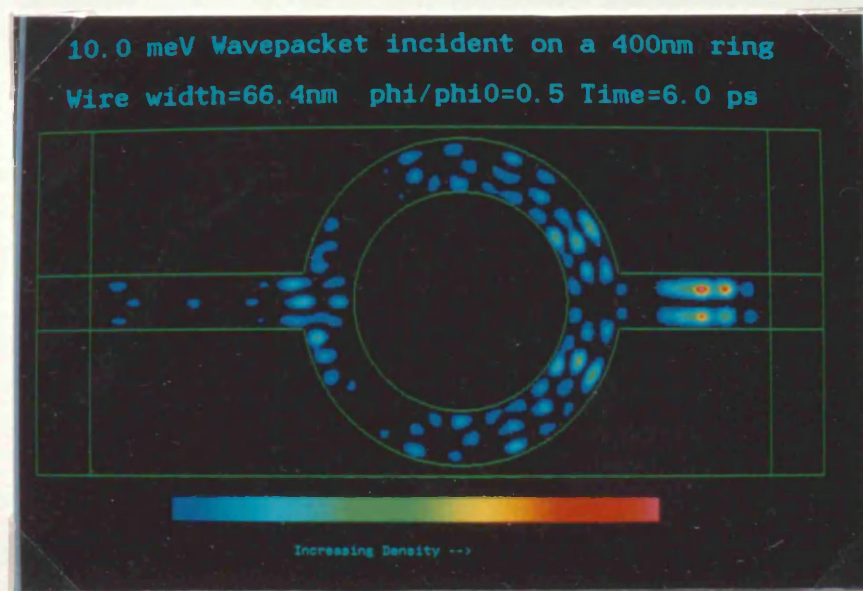


Fig 8.48

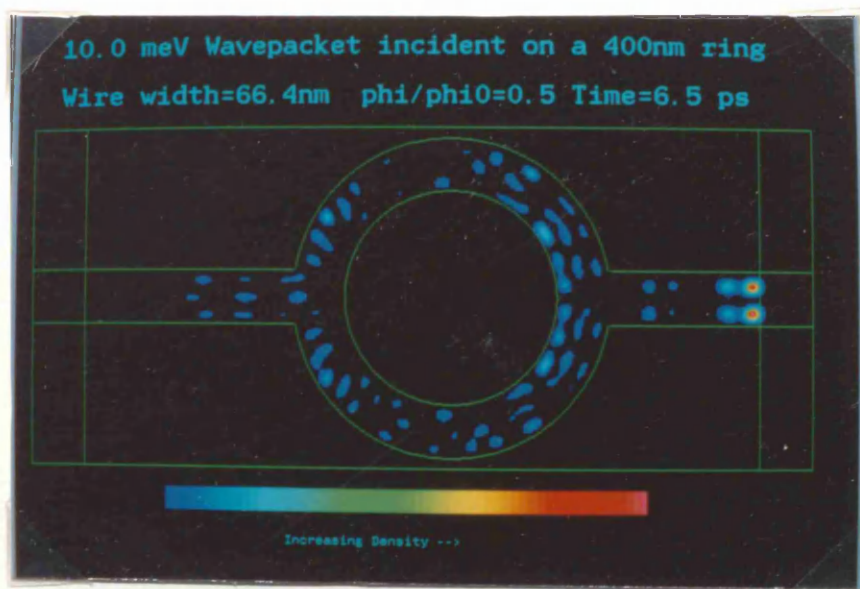


Fig 8.49

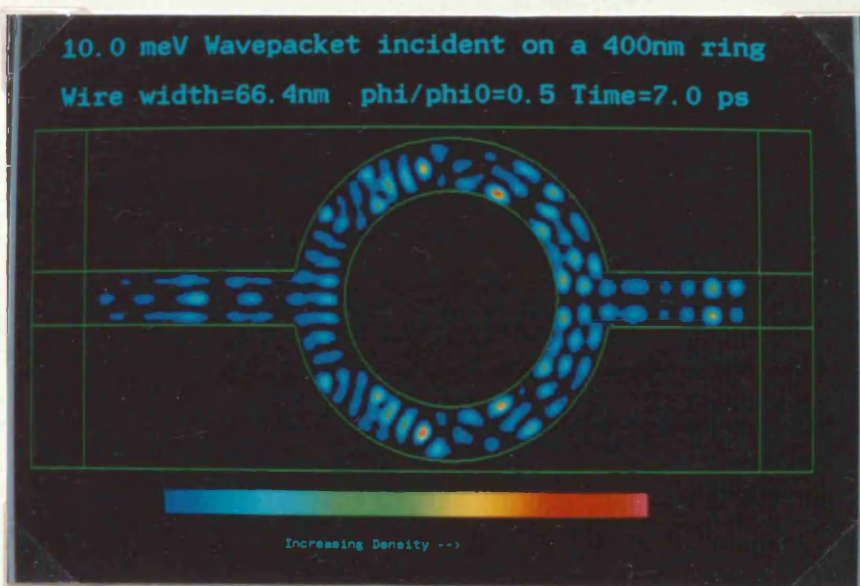


Fig 8.50

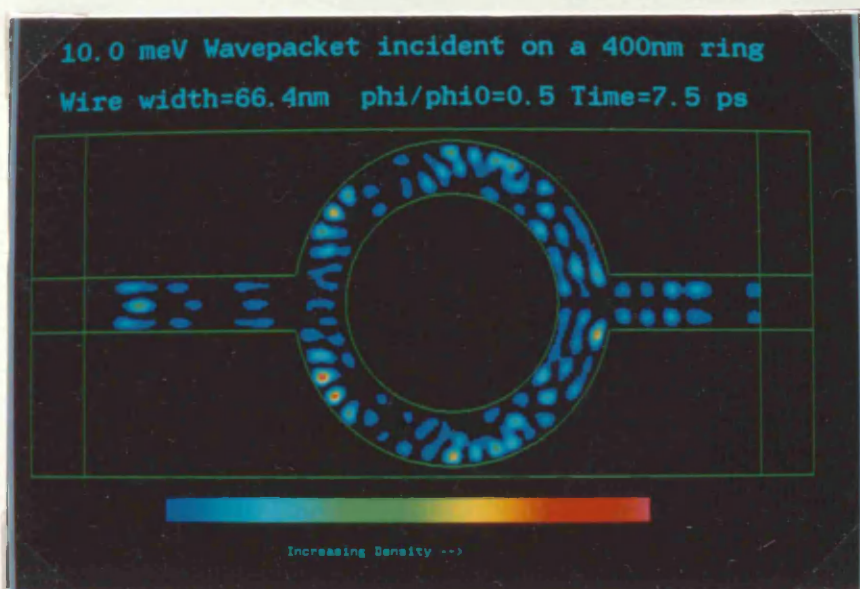


Fig 8.51

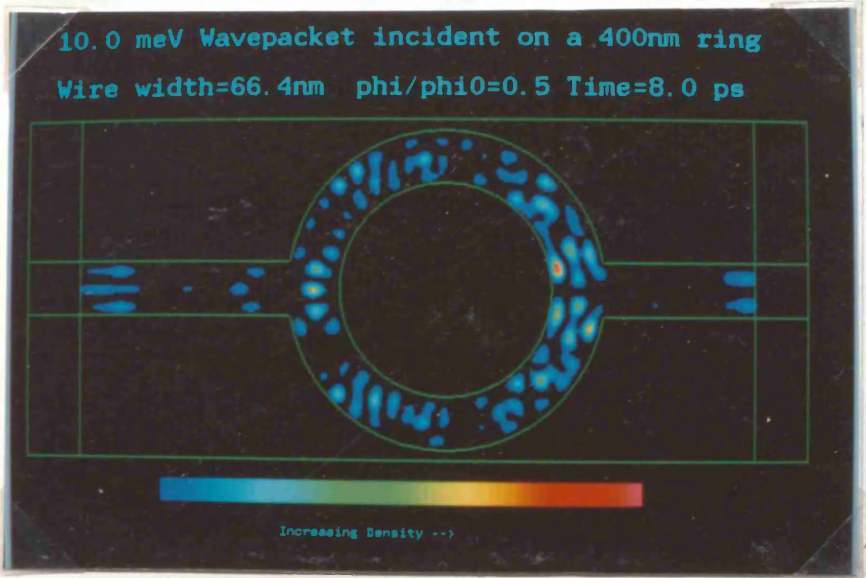


Fig 8.52

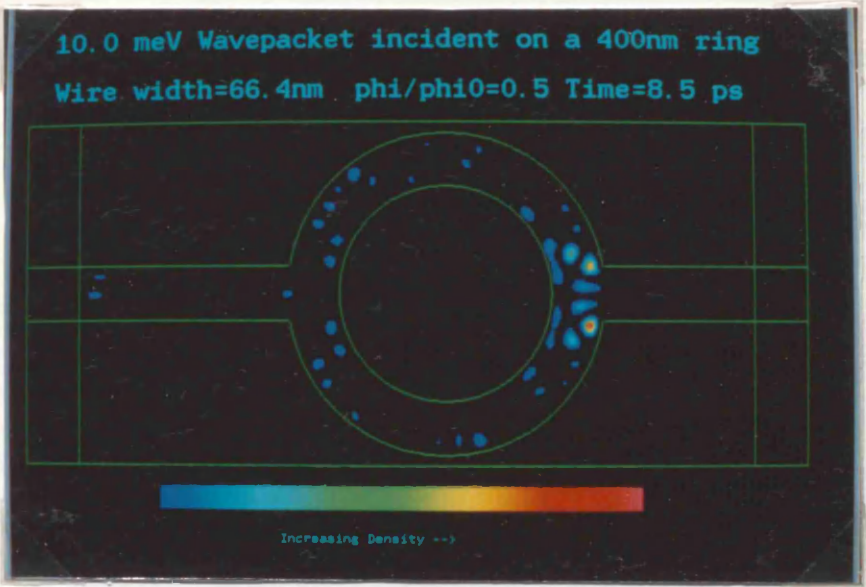


Fig 8.53

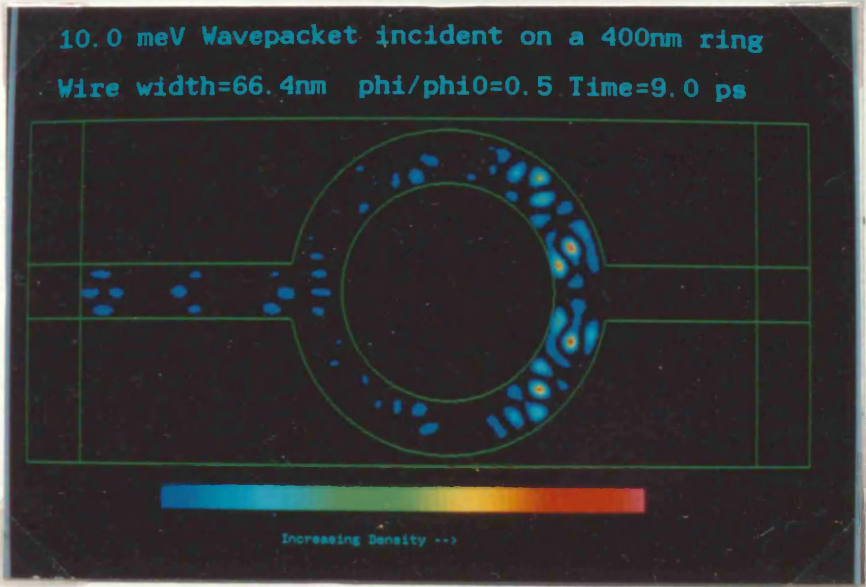


Fig 8.54

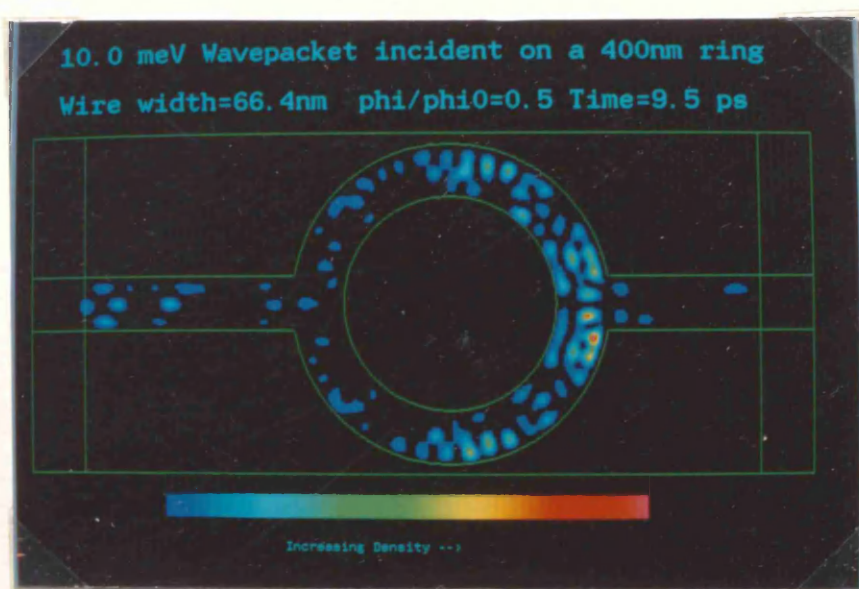
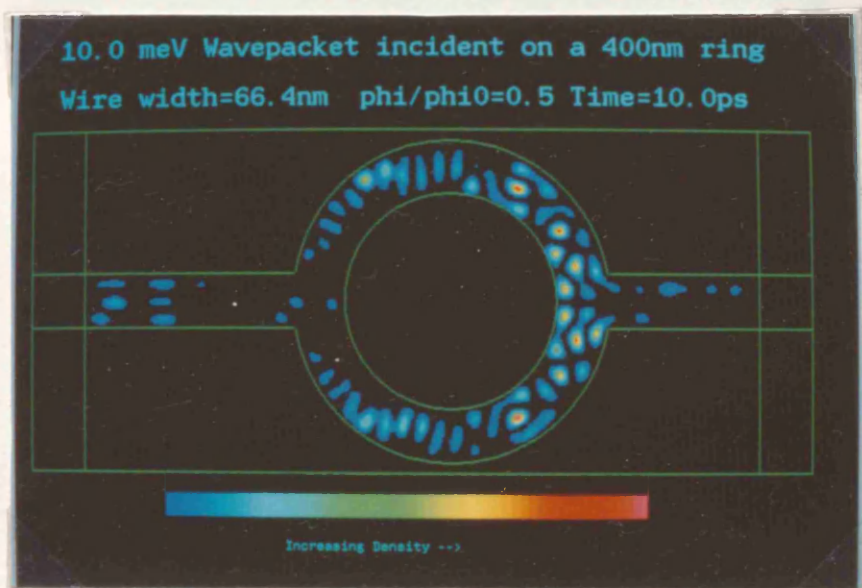


Fig 8.55



Figures 8.56 to 8.58: 3D representations of the wavefunction for threaded fluxes of 0, $0.5(h/e)$ and h/e respectively.(uniform field)

Fig 8.56

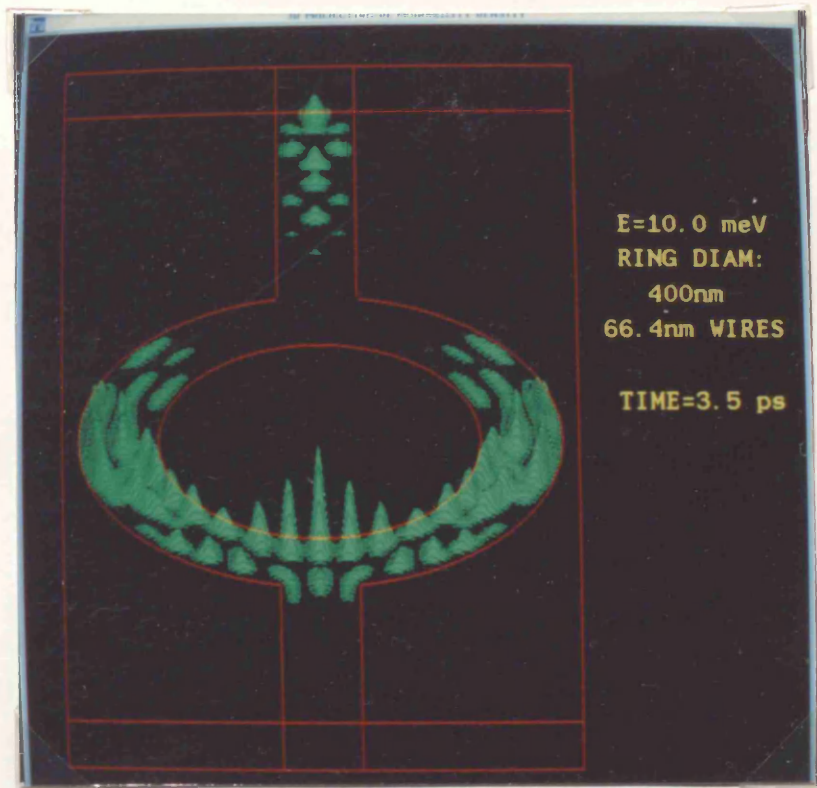


Fig 8.57

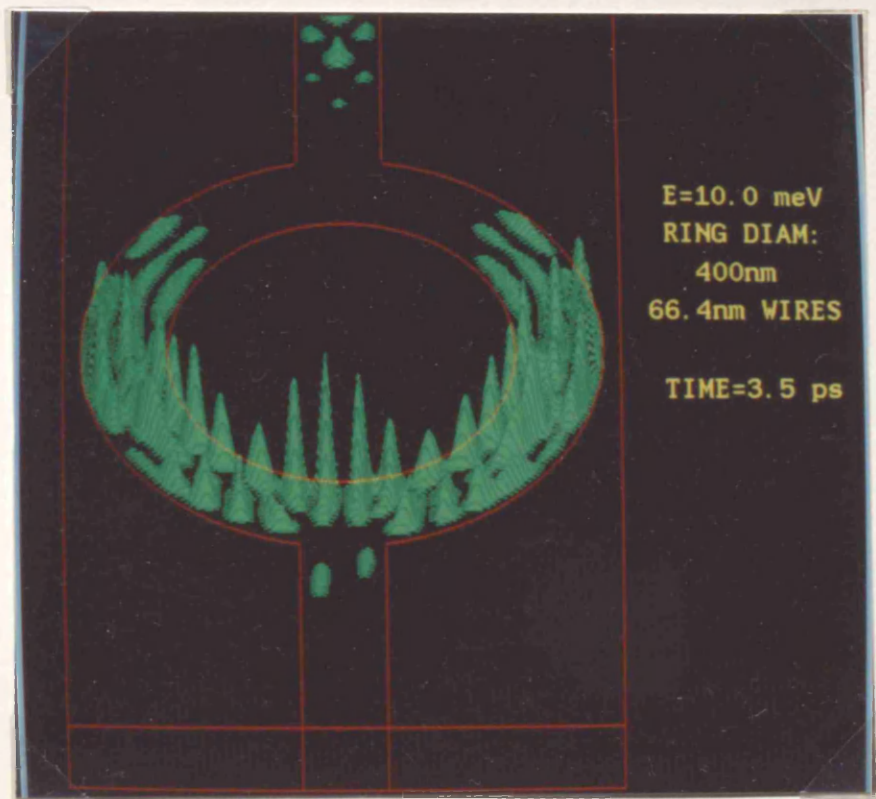
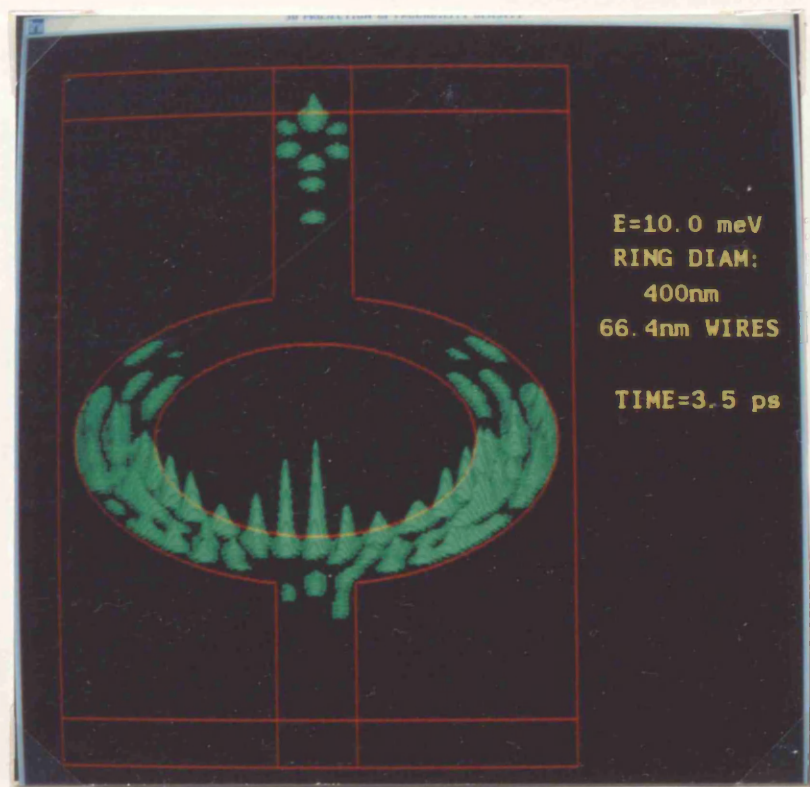


Fig 8.58



Figures 8.59 and 8.60: 3D representations of the wavefunction for threaded fluxes of 0 and 0.5(h/e) respectively.(uniform field , time from launch=4.5ps)

Fig 8.59



Fig 8.60



8.4.2 Multimode structure.

The second feature of importance in the large-scale simulations was the emergence of a multimode structure in the wavefunction. This is caused by the transfer of forward momentum in the initial wavepacket to transverse modes across the wire as the wavepacket collides with the inside wall of the ring at the input junction.

On impacting the ring, the wavefunction is seen to split into three modes across the wire. An analysis of the modes present half way round the upper arm of the ring was performed by assuming that the wavefunction across the wire could be expanded in terms of the particle-in-a-box eigenfunctions for a potential well with infinitely high walls. The expansion coefficients were then found by projection of the computed wavefunction across a particular line onto this basis set. The results confirmed the third eigenfunction to be the dominant mode, typically accounting for 80-90% of the wavefunction present in the arms. Of the remainder, the first sub-band wavefunction was generally the largest. The higher fourth, fifth, and sixth sub-band modes had very low amplitudes.

The initial forward energy component of 10meV plus the transverse energy component of 1.27meV is slightly less than the transverse energy of 11.46meV of the third sub-band. However, assuming the initial wavepacket does not significantly broaden before impacting the ring, the spread in energies contained within the wavepacket mean that one can expect about 45% of the initial wavepacket to have sufficient energy to occupy the third sub-band. In the time-dependent situation, the other sub-bands will mix with the third sub-band eigenfunction to give a time-dependent oscillation of the probability density from side to side across the wire. When viewed as a time sequence, the highest peaks in the transverse direction suggest a classical trajectory of a point particle bouncing around the ring (apart from near the output junction where interference effects with the wavefunction in the other arm become important).

Figures 8.36 to 8.55 show the same wavepacket incident on the same ring as described above when a uniform magnetic field is applied over the entire ring. The magnetic field was chosen such that the area inside the ring, defined by the average radius, enclosed half an h/e flux quantum. The sequence of frames is initially virtually indistinguishable from the zero-flux sequence; the difference only becomes obvious when the two partial wavepackets start to interfere, when a trough instead of a peak

appears at the centre of the output wire. This is the same type of behaviour that was seen in the preliminary studies which again demonstrates the AB effect in this ring structure and also verifies that the effects of the flux-induced phase change are only apparent when the partial wavepackets interfere.

One of the main results from this simulation is that it reveals the multi-mode structure in the probability density as being responsible for the over-estimate of the magneto-transmission modulation depth as based on 1D models. Even when there is a minimum in the wavefunction present at the centre of the output wire, the side-lobes of the interference pattern can still propagate out of the ring, leading to a non-zero transmission.

A picture now emerges in which the experimental situation can be viewed as being intermediate between two extreme cases. Imagine, for example, that a phosphor screen could somehow be placed at the exit of the ring. A particular phosphor would glow each time an electron impacted it, and the total current drawn-off on a back-plate and measured. Thus, at one extreme there is the perfect one-dimensional case in which perfect current modulation as a function of magnetic flux can be achieved, but only a minimal knowledge of the resulting interference pattern can be gained (one phosphor either on or off). At the other extreme, the wire is made infinitely wide. In this case a complete knowledge of the interference pattern may be gained but no modulation in current will be detected across the plane of the screen. This is because the magnetic flux only creates a re-distribution in the wavefunction, so that no change in the current should be observed if the measurement domain (the wire) is made large enough.

Although the complete transmission process can only be fully described by the two-dimensional model, the above observations offer an interpretation in terms of quasi-1D transport in which four stages can be identified.

- 1) The wavepacket, (initially in the first sub-band) moves into the input-junction area and encounters a widening of the wire. The wavepacket spreads-out and reforms into three transverse modes as it begins to propagate around the arms.

- 2) The wavepacket propagates round the ring, suffering additional interference in the longitudinal direction due to both reflection from the output junction and the other partial wavepacket.

3) The reverse process to stage (1) occurs. The wavefunction, as it nears the output junction, can be thought of as forming transverse modes across a wide wire.

4) The transverse modes induced across the output-side of the ring can then excite modes across the output wire, and some of the wavefunction propagates out of the ring in these excited modes.

Figures 8.56 to 8.58 show the position probability density for the ring described above for threaded flux values of 0.0, $1/2(h/e)$ and (h/e) respectively plotted in a 3D perspective looking from the output direction of the ring at an angle of 55 degrees from the normal. Figures 8.56 and 8.57 again clarify the effect of a threaded flux of $1/2(h/e)$ in changing the central maximum which is centred on the output wire to a minimum. In figure 8.57, the two side lobes either side of the centre-line of the output wire can be clearly seen and are free to propagate out of the ring, which it is proposed, are responsible for the loss of modulation depth in the magneto-transmission characteristics. The slight asymmetry in the interference pattern arises from the Lorentzian force acting on the wavepacket caused by the (uniform) magnetic field penetrating the wires in addition to the interior of the ring. Figure 8.58 shows the probability distribution from a simulation in which the flux was increased to the full flux quantum h/e . General theory predicts that if the flux were solely confined to the interior of the ring, figures 8.56 and 8.58 representing the cases of 0 and h/e enclosed flux values respectively, should be identical. It can be seen that the two plots match very closely, but the plot with the magnetic field applied is again slightly asymmetric with respect to the output wire as a result of the Lorentzian force. Figures 8.59 and 8.60 show the probability distribution for the wavefunctions of figures 8.56 and 8.57 one picosecond later and demonstrate how the central peak (or two secondary fringes in the case of figure 8.60) propagate out of the ring.

From figure 8.58, the distance that the fringes have been shifted was found to be approximately 3nm in the clockwise direction. To obtain an estimate of the behaviour of the overall envelope shift at higher magnetic fields, a relationship between the spatial shift and field is required. An analysis of the effect of a uniform field on the interference pattern in the biprism experiment [Olariu and Popescu 1985] showed that for small deflections, the shift varied linearly as the magnetic field. Assuming this relationship, the constant of proportionality, derived from measurements on the 3D

plots is about 66nm per Tesla. It is proposed that the magnetic field will start to appreciably damp the amplitude of the oscillations when the centre of the interference envelope has been shifted so that it is no longer co-incident with the output wire. Substituting a value of half the wire width for the envelope shift into the derived relation, yields a value of 0.5T for the corresponding flux for this ring. Thus, it is predicted that the AB oscillations should start to disappear at a field of about 0.5T. This result is consistent with the experimental observation that the oscillation amplitude starts to decrease in the range 0.5-1.0T and eventually disappear at fields between 1.0 and 1.3T [Ford *et al* 1988]. These comparisons should, however, be treated with caution as the oscillations in the experimental data are rather intermittent.

In addition to the standing-wave structure due to interference between the two partial wavepackets as they meet at the exit of the ring there is also some interference due to the reflection of each of the partial waves with themselves within their respective arms. The reflection of the partial waves from the output junction back into the ring was predicted in the one-dimensional analysis in chapter 5, and was found to give rise to higher harmonics in the magneto-transmission characteristic.

In order to isolate the effect the reflection from the output junction in the two-dimensional model, an absorbing region, similar to that used for the contacts, was placed in the lower arm of the ring. This absorbed any wavefunction in the middle of the lower arm, leaving only the interference effects due to one of the partial wavepackets. Figures 8.61 to 8.80 show the results of these calculations.

The incident wavepacket is seen to split into two partial wavepackets, each having three transverse modes, similar to before, showing that the splitting of the wavefunction into three transverse modes is only dependent on the local input junction geometry. After 3.5ps from launch, the wavepacket in the upper, unblocked arm, begins to acquire separate peaks in the direction of motion, which is different to the uniform broadening of a single peak one would expect from a free wavepacket. The appearance of peaks is an indication that interference is occurring and since the other partial wave has been blocked, the interference can only be arising from the upper partial wavepacket interfering with itself. This supports the hypothesis based on the 1D model that there is inherent reflection present at the output junction.

Figures 8.61 to 8.80: Propagation of a wavepacket through a 400nm diameter ring, showing the effect of interference along just one arm.

Fig 8.61

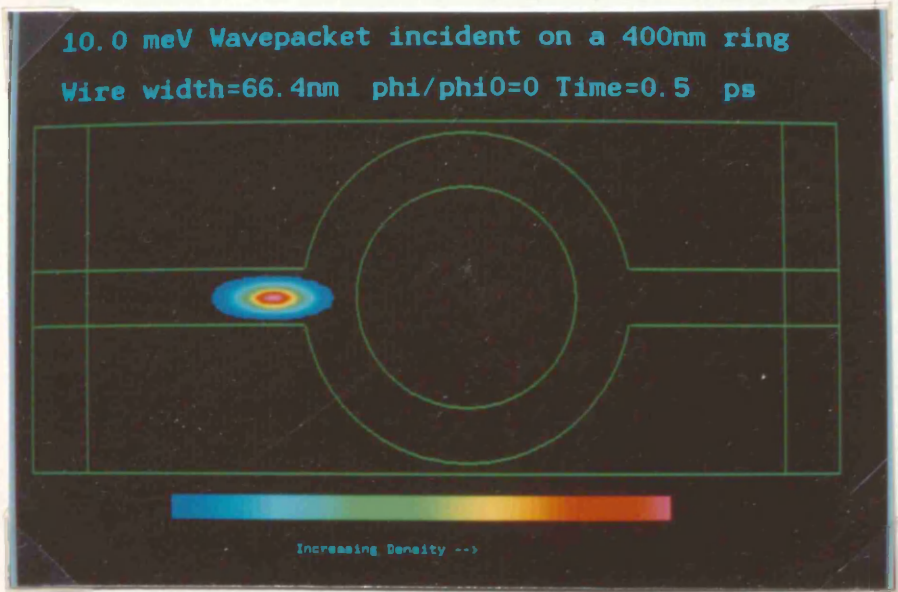


Fig 8.62

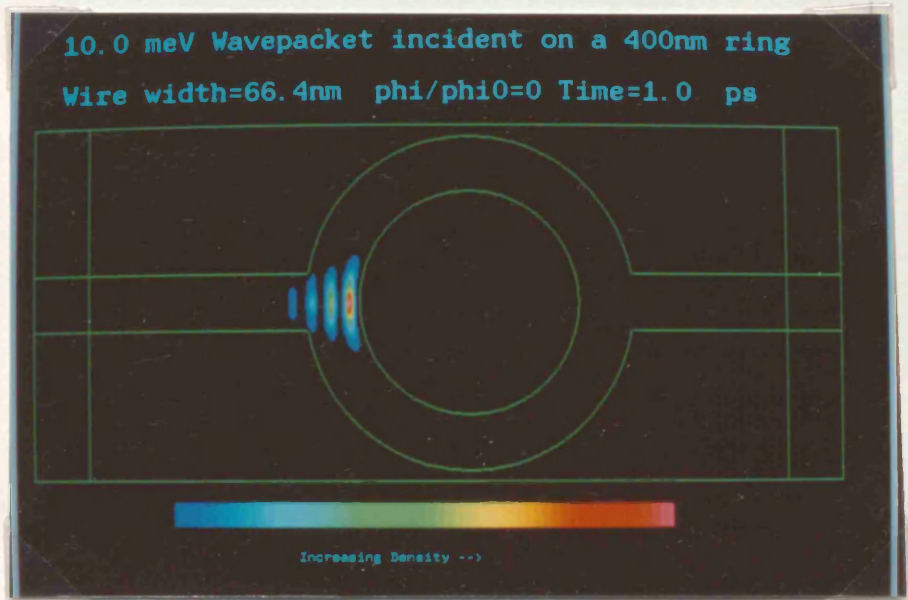


Fig 8.63

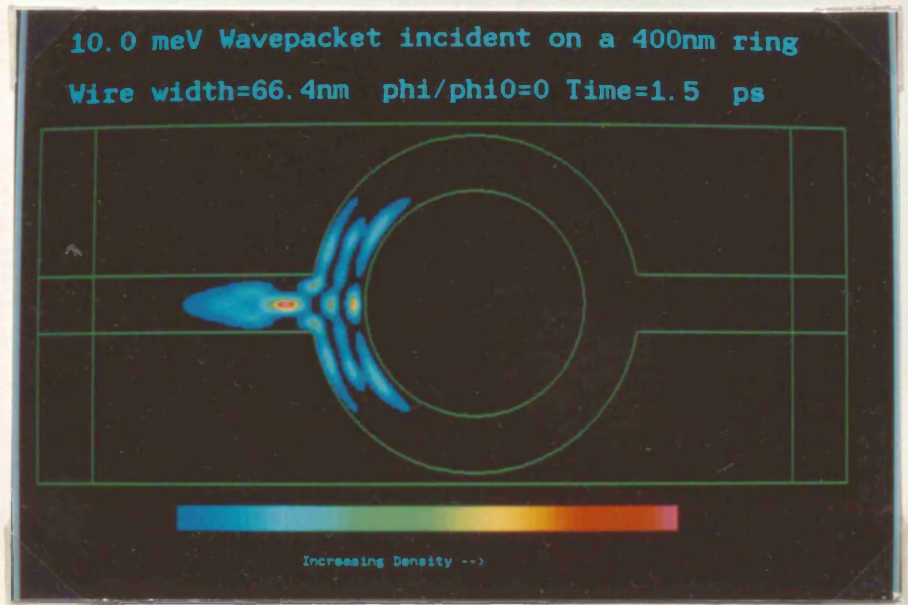


Fig 8.64

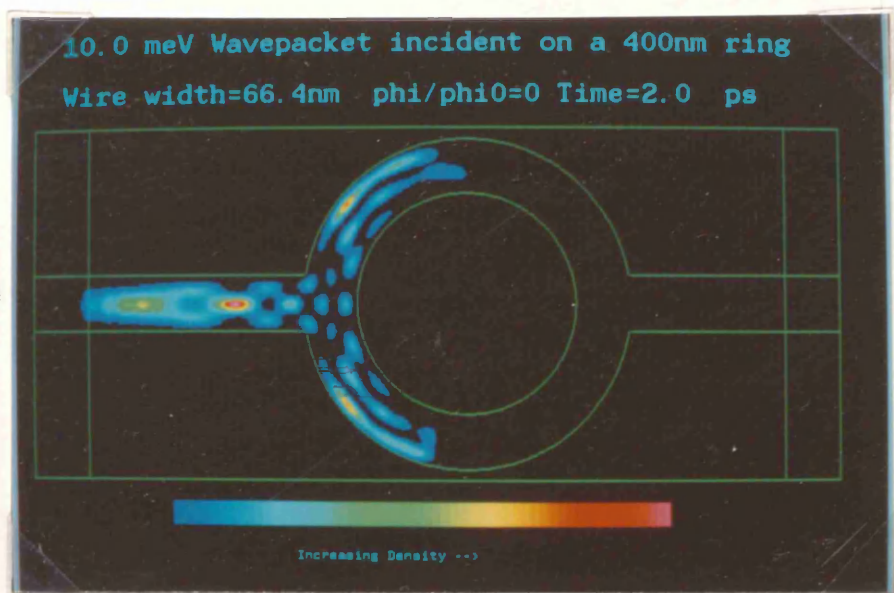


Fig 8.65

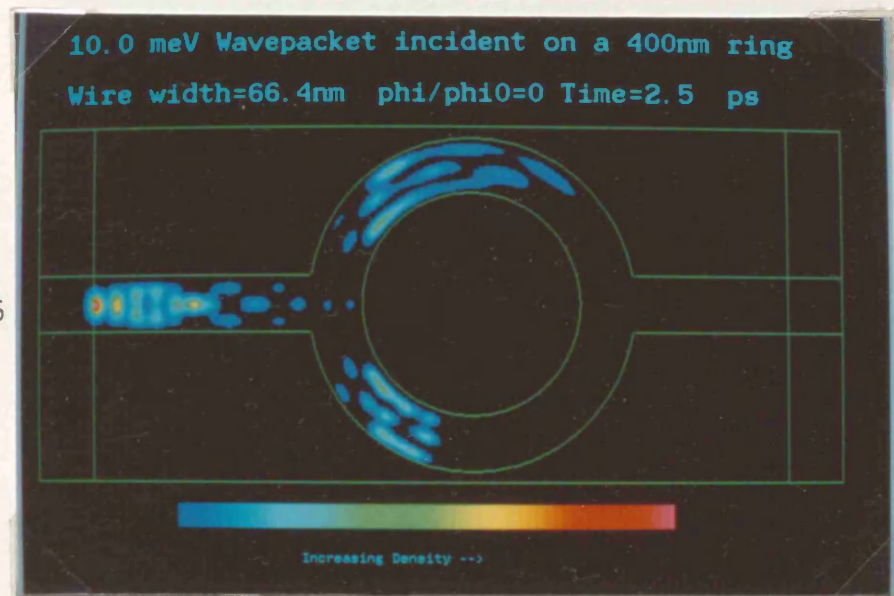


Fig 8.66

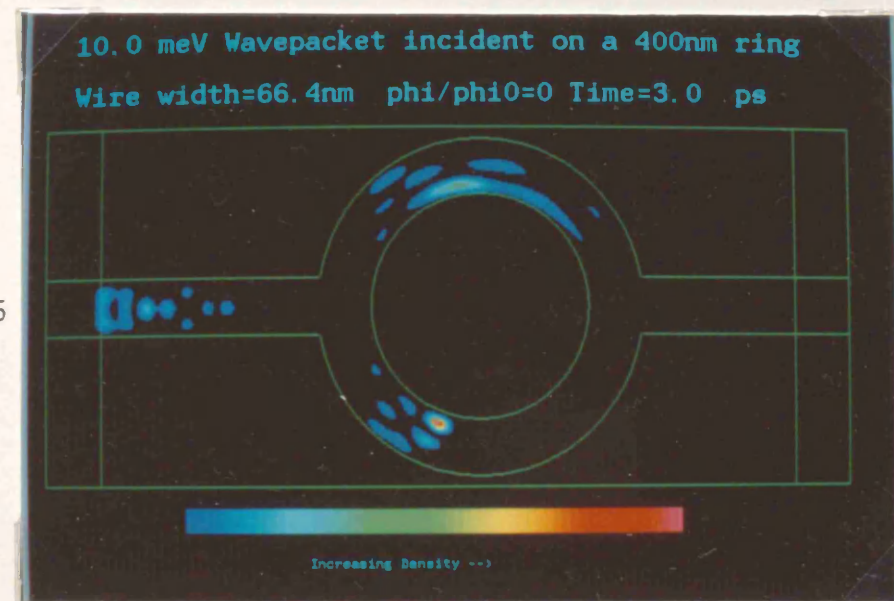


Fig 8.67

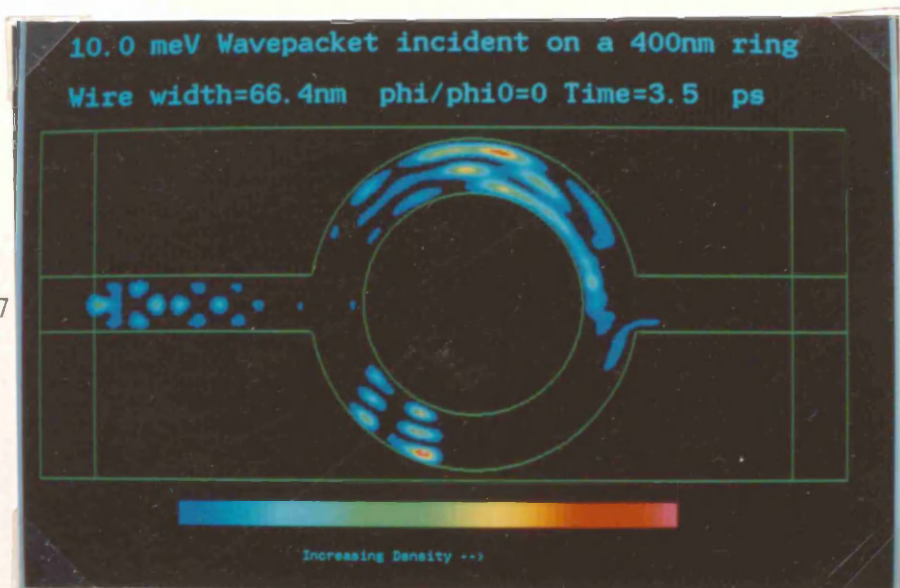


Fig 8.68

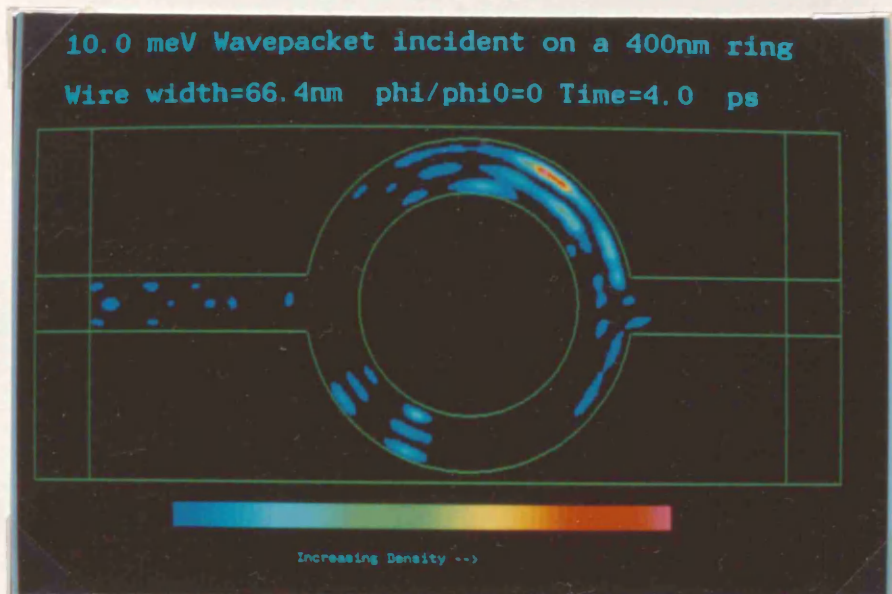


Fig 8.69

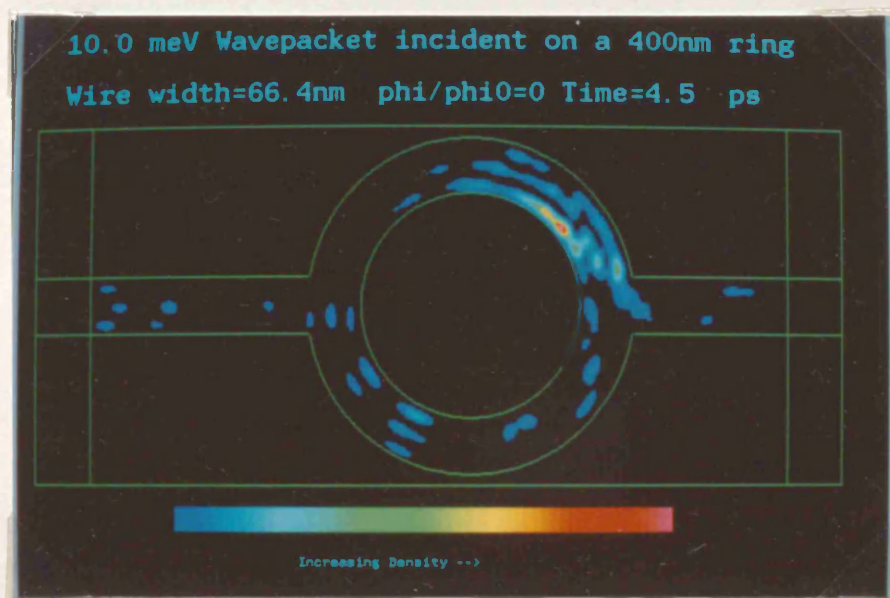


Fig 8.70

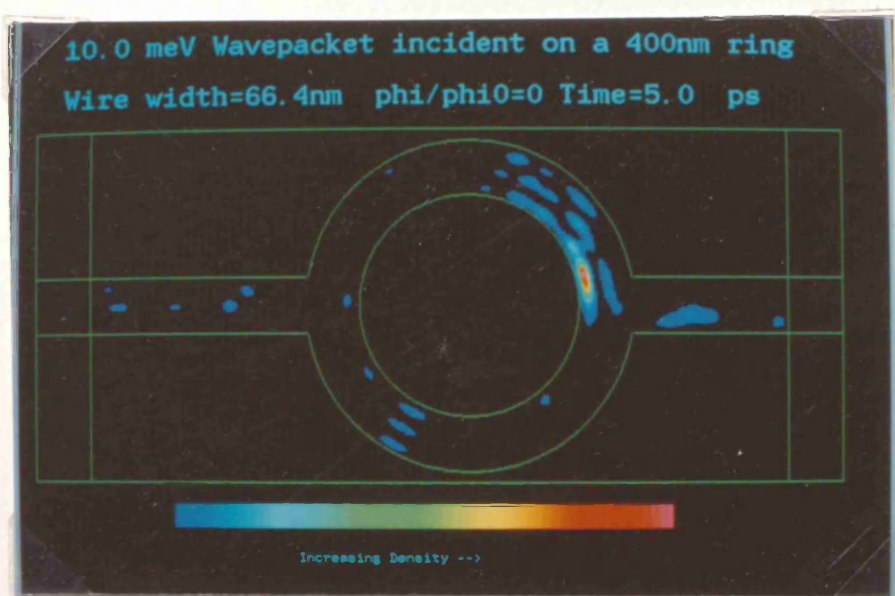


Fig 8.71

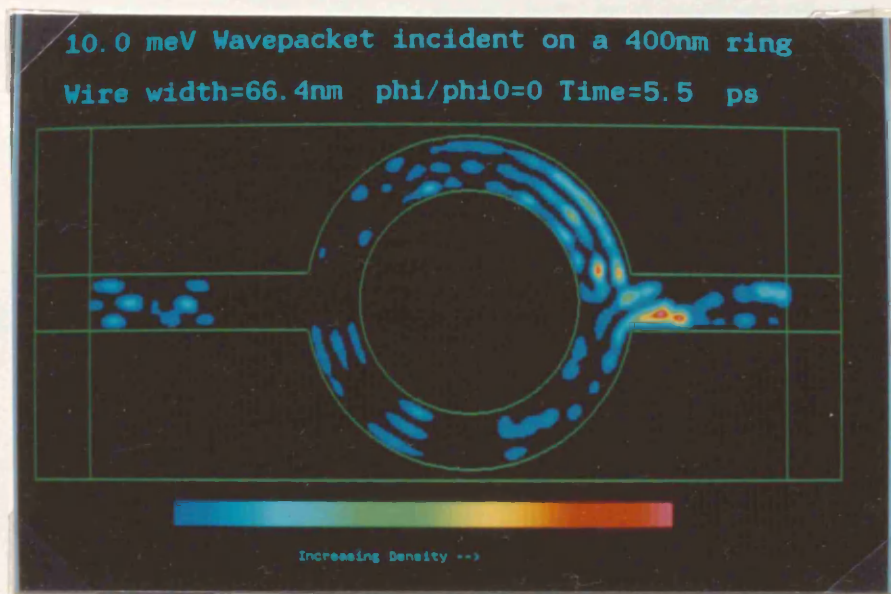


Fig 8.72

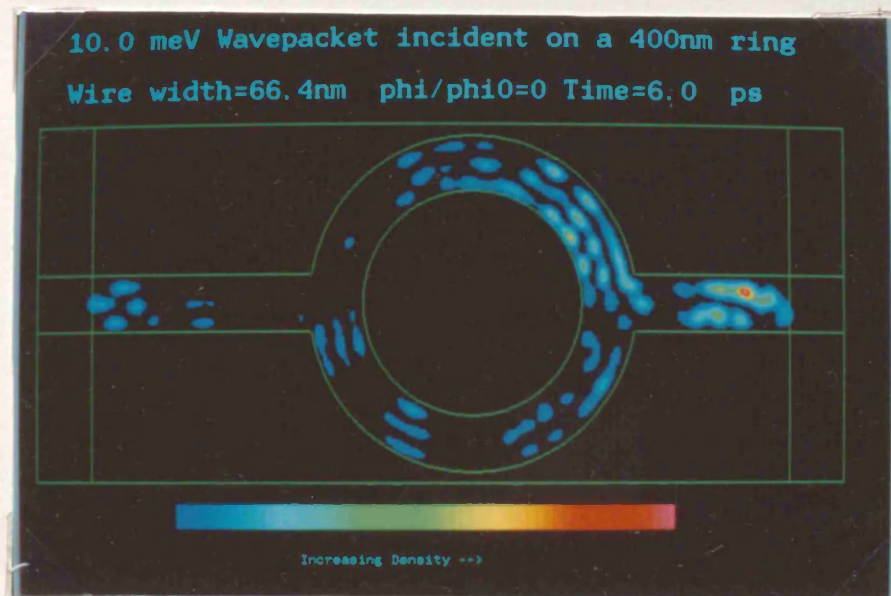


Fig 8.73

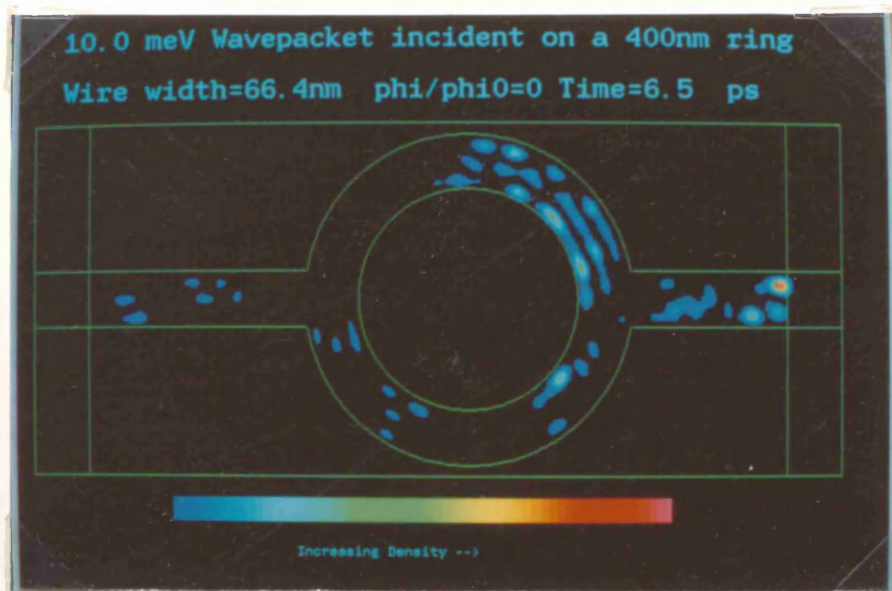


Fig 8.74

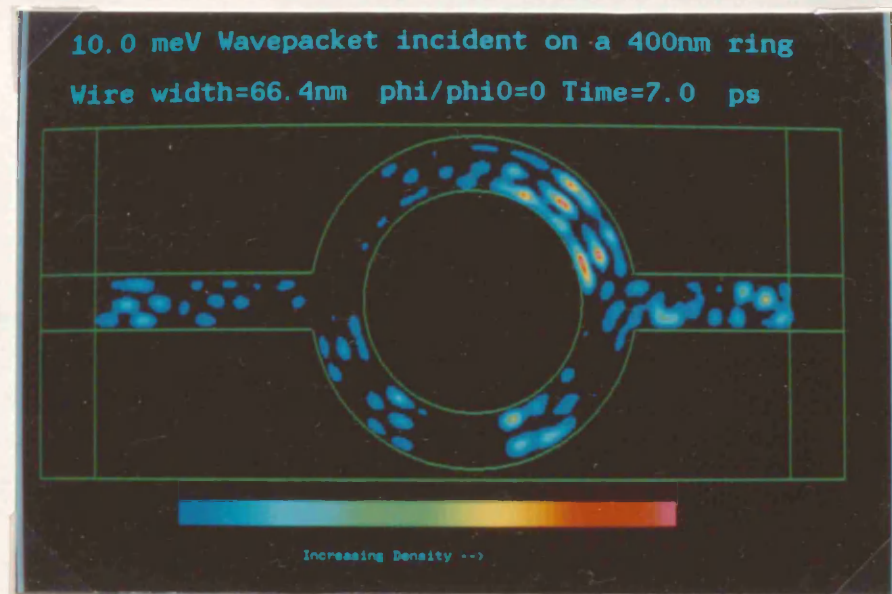


Fig 8.75

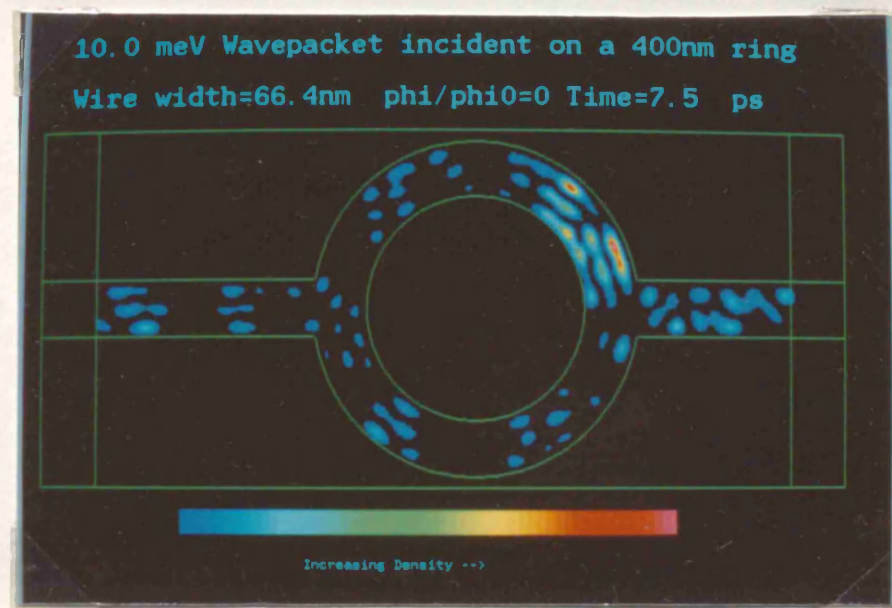


Fig 8.76

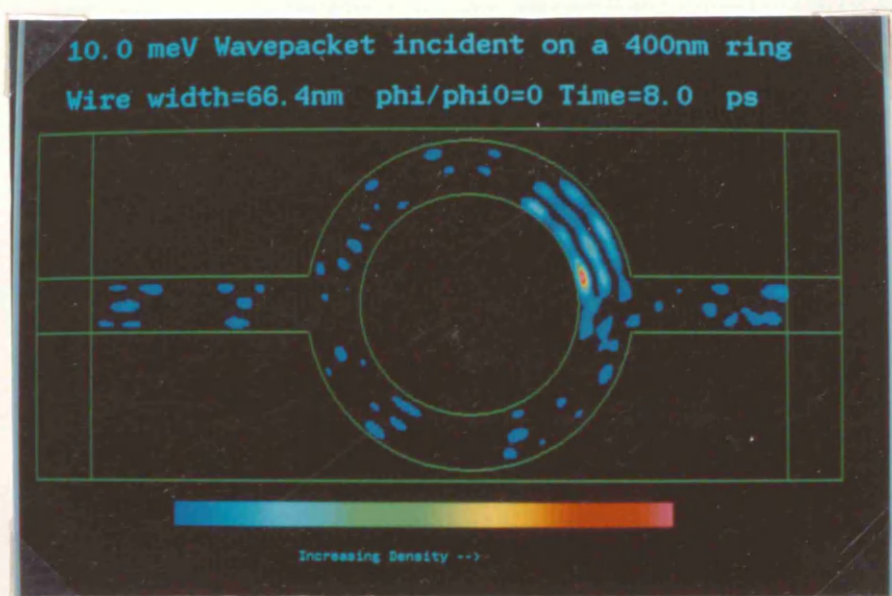


Fig 8.77

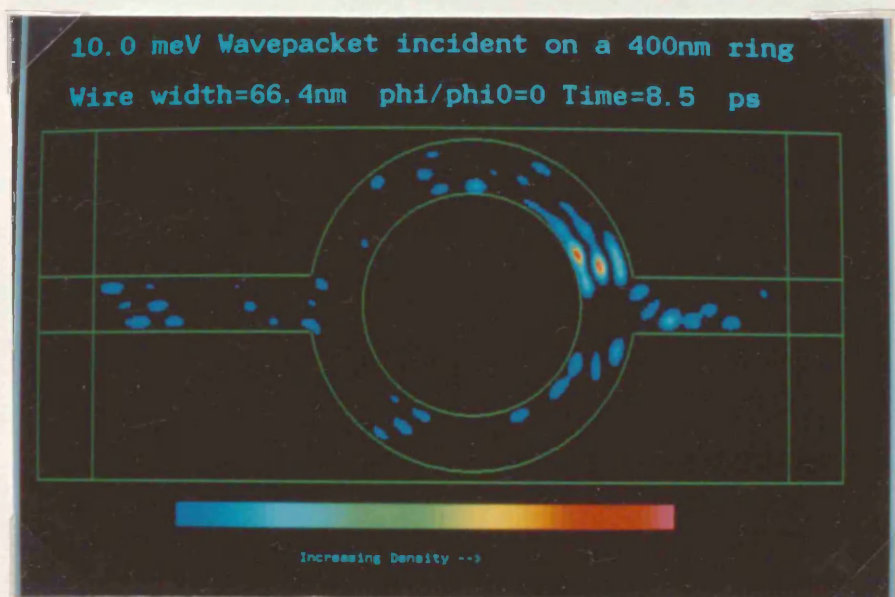


Fig 8.78

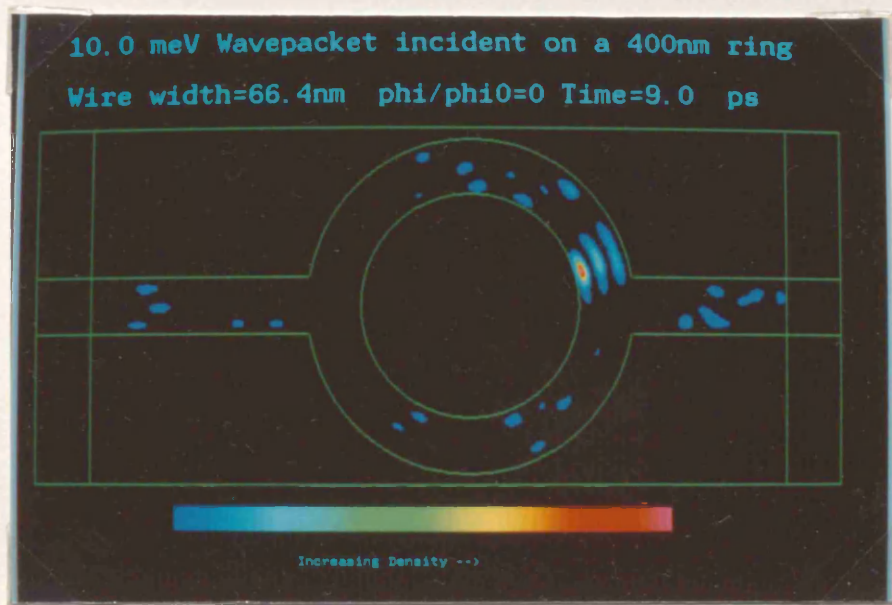


Fig 8.79

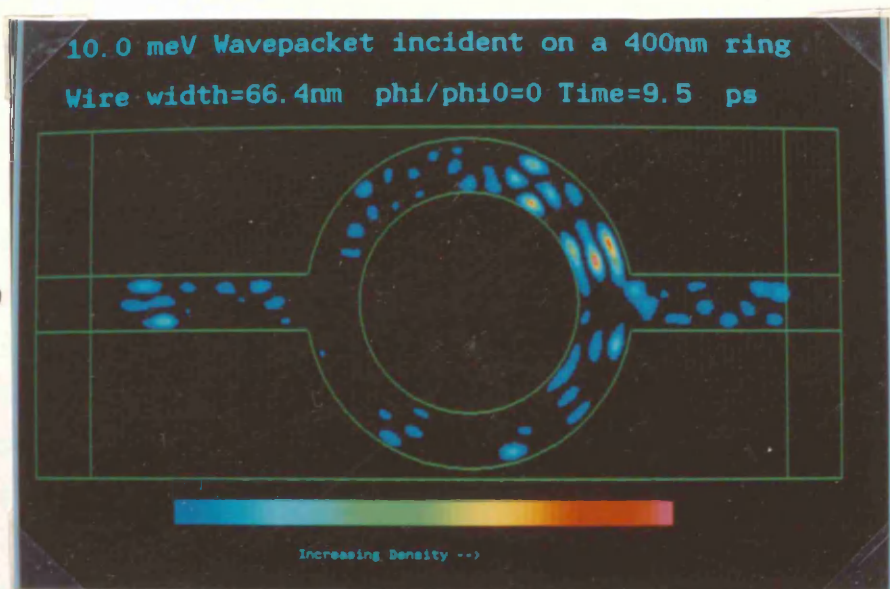
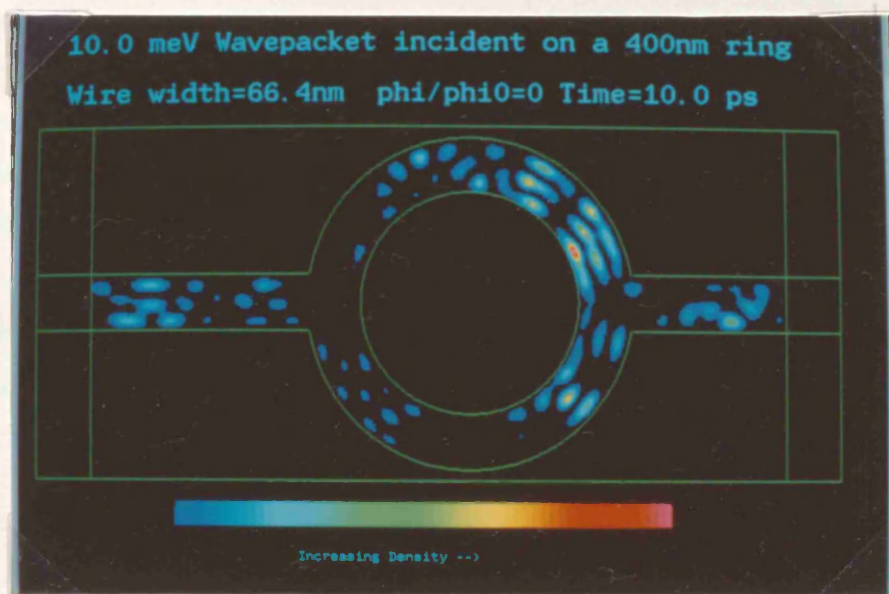


Fig 8.80



8.4.3 Transit times.

Like the double-barrier problem, there are several ways in which the determination of transit time can be approached. In this study the problem was treated in two different ways. The first method was to measure the position of the wavefront approximately from the contour plots as it propagated around the ring. The second method was to calculate the transmitted charge within the simulation as a function of time. The first method, although approximate, does provide information on how the wavepacket propagates around the ring before it contributes to the probability current measured at the contacts.

In both methods, the measured positions and currents were compared to those of the same initial wavepacket travelling an equivalent distance down a wire of the same width in a straight line. Since a free wavepacket travelling in a straight and uniform wire experiences no mechanism whereby it can transfer momentum to higher sub-bands, the problem can be reduced to the solution of the free wavepacket problem in one dimension.

For a normalised Gaussian wavepacket of initial average wavevector K_0 and position standard deviation Δx , the solution to the time-dependent problem, obtained by expanding $\psi(x,0)$ into a set of plane-wave states is:-

$$\psi(x,t) = (\pi\Delta x^2)^{-1/4} (1+i\hbar t/(m^*\Delta x^2))^{-1/2} \exp(-0.5K_0^2\Delta x^2 + iK_0x_0) \\ \times \exp((K_0^2\Delta x^4 - (x-x_0)^2 + 2i(x-x_0)K_0\Delta x^2)/(2\Delta x^2 (1+i\hbar t/(m^*\Delta x^2))))$$

(Eqn 8.1)

The probability current was calculated at a point x , which was arranged to be the same distance away from x_0 as the centre of the initial wavepacket in the simulations was from the current measurement point. In the simulation the transmitted current was measured 3.32nm before the right hand side contact and the distance from the centre of the initial wavepacket to the measurement point was taken as the distance along the centre of the wire around any one of the arms.

The results derived from measuring the propagation of the wavefronts were compared to the free wavepacket by calculating the probability density. The position where the probability density was 2/60ths its maximum value at time t , was plotted alongside the results from the simulation as a function of time.

Figure 8.81 shows a plot of the distance of the wavefront of a 10meV wavepacket from its initial position as a function of time when the magnetic field is zero. The straight dotted line is the analytic result for the same initial wavepacket in one dimension and has a slope of about 270×10^3 m/s which is slightly higher than that of a 10meV plane wave due to wavepacket spreading.

It is seen that the simulated motion is in good agreement with the analytic result up to about 0.5ps, which is when the wavepacket begins to impact the ring. From about 0.5ps to about 2ps a decrease in velocity of the wavefront was observed. It was thought that the transfer of the wavefunction to higher sub-bands might cause a decrease in forward momentum but the amount by which it decreased was not as great as was expected. Figure 8.82 shows the wavefront position as a function of time of a 1.27meV wavepacket in which the propagation of the wavepacket was single-mode. This graph is qualitatively the same as the graph for the 10meV wavepacket in that the wavefront propagation is in reasonable agreement with the analytic result up to the point the partial waves begin to interfere at the output junction. This suggests that the wavefront propagation for the 10meV wavepacket is largely due to the single-mode component of the wavepacket.

Both curves start to diverge from the analytic result when the partial wavepackets begin to interfere; the curve for the 10meV wavepacket diverging the most rapidly. One reason for this is that the plotted probability density is scaled to the maximum value of the density in the ring. The free wavepacket continues to decrease in absolute peak amplitude and broadens. The partial wavepackets in the simulation are subject to an increase in peak density in the interference fringes when they meet at the output. This means that the rest of the wavefunction is scaled to these peaks which has the effect of defining the emerging wavefront position as being nearer to the output junction.

Figures 8.83 and 8.84 show the transmitted charge as a function of time for a 10meV and 1.27meV wavepacket respectively. In both cases the analytic result for a free wavepacket is shown in a dotted line and is normalised to the asymptotic transmission coefficient derived from $\text{t}_{\text{th}} \lambda$ simulations.

The transit time, defined here as the time taken for the wavepacket to achieve 90% of its asymptotic transmission is 11.8ps compared to a free wavepacket transit time of 2.2ps. The transmitted charge approaches the asymptotic value in an approximately exponential fashion, which supports the view that the wavepacket undergoes many internal reflections before escaping. This interpretation is later used in chapter 9 to explain the shape of the Fourier transform spectra of the magneto-transmission data. Figure 8.84 showing the transit time of a 1.27meV wavepacket shows a similar type of behaviour to the 10meV wavepacket apart from the appearance of distinct "steps" in the transmitted charge as a function of time. This provides further evidence that the wavefunction trapped inside the ring is escaping in bursts as it oscillates backwards and forwards from the input to output junctions.

In conclusion it is commented that when several modes are possible across the wire, the time taken for the whole wavepacket to propagate around the ring will be longer than the time taken for the same wavepacket in the first sub-band to travel the equivalent length in a straight wire. This is because the wavepacket can transfer to higher-order transverse modes when it impacts the centre of the ring, due to the increase in the transverse component of the energy and the consequent reduction in the forward component.

The collision of the wavepacket with the interior of the ring can be viewed as a special case of an elastic scattering event of the sort normally encountered with impurity potentials. The interior of the ring can thus be regarded as a single giant impurity potential, its effect being to reduce the mobility of the wavepacket.

In a straight uniform wire, there is no mechanism whereby any transfer of momentum to higher sub-bands can occur if the motion is again considered to be ballistic within the wire. If an electric field is applied to a length of wire, over which transport is ballistic, the wavepacket will increase its forward momentum without scattering to higher sub-bands. In the case of the initial condition being a mixture of transverse modes, one can treat each sub-band independently so that the total ballistic resistance is the sum of the resistances for each of the sub-bands connected in parallel [Wharam *et al* 1988].

Figure 8.81

Approximate transit time of a 10meV wavepacket through a 4000 Angs. AB ring.

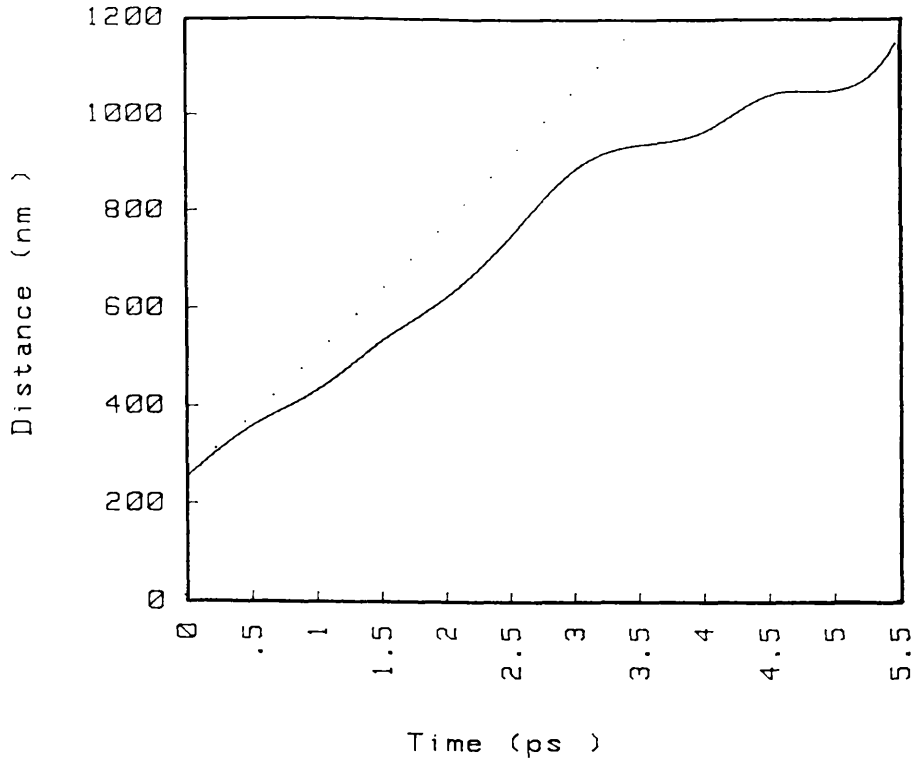


Figure 8.82

Approximate transit time of a 1.27meV wavepacket

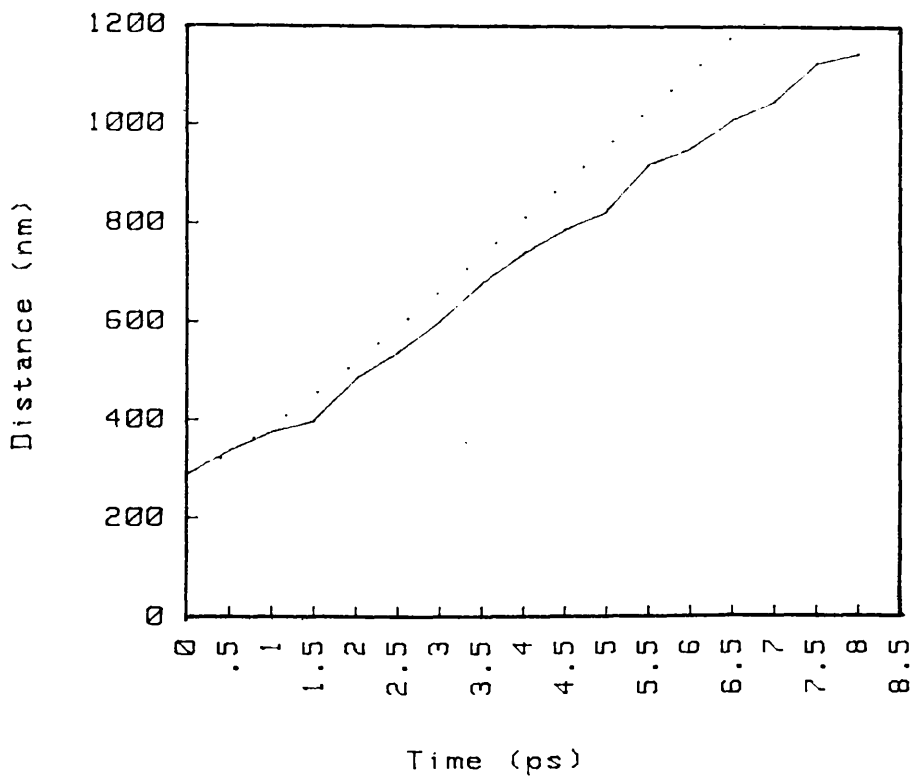


Figure 8.83

Transit time of a 10meV wavepacket through a 400nm ring

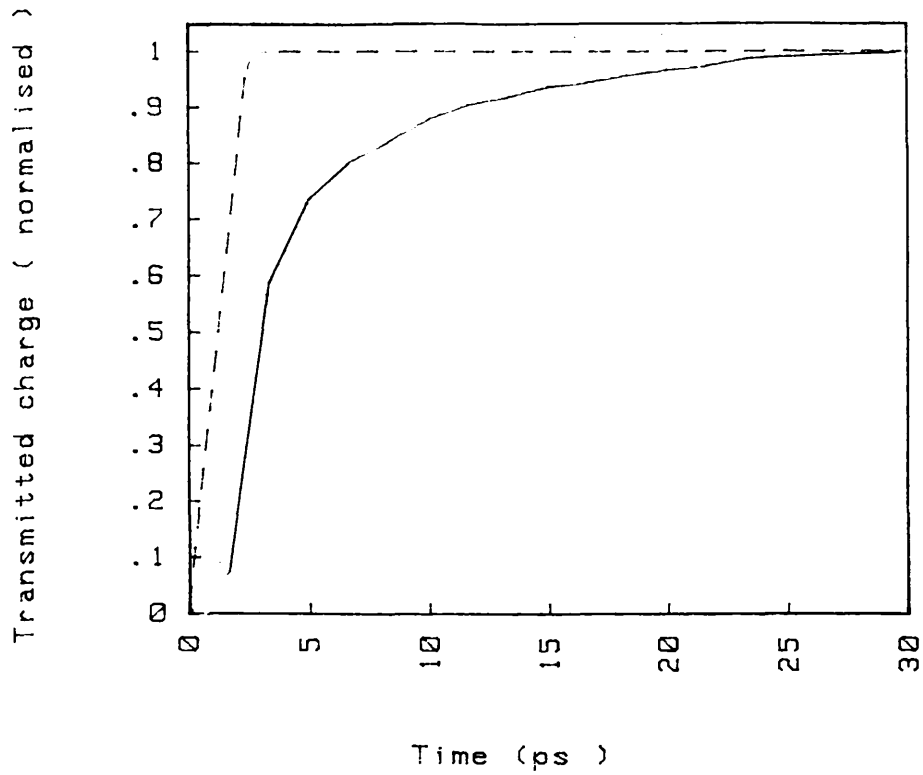
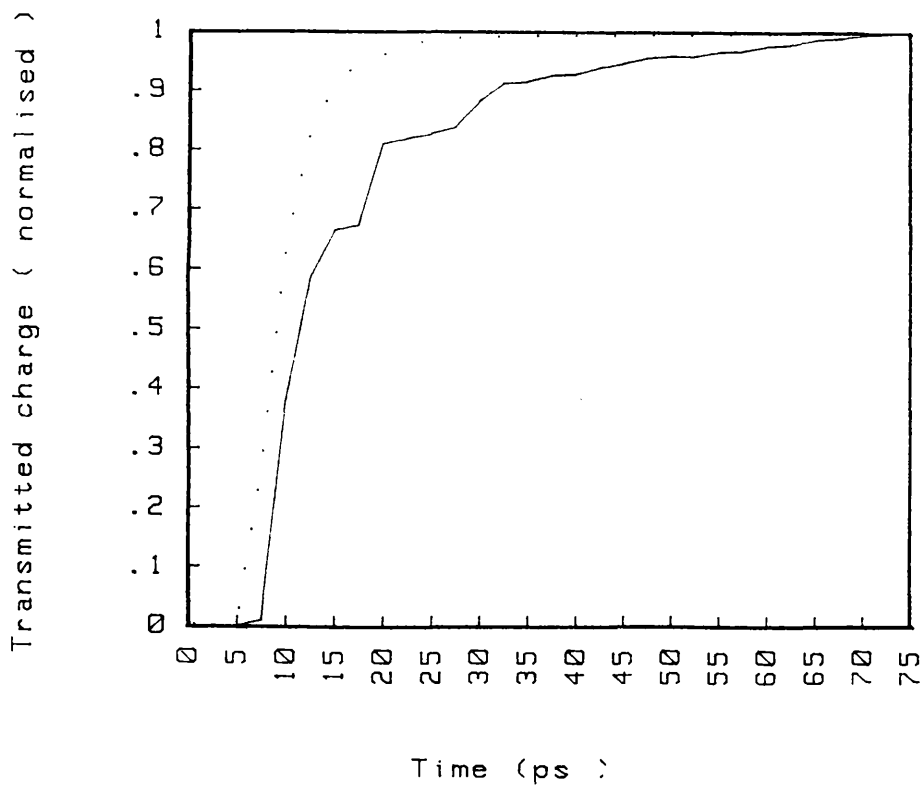


Figure 8.84

Transit time of a 1.27meV wavepacket through a 400nm ring



8.5 Summary

This chapter has presented the results of two-dimensional time-dependent simulations of wavepackets incident on AB rings.

Preliminary results from a very small ring structure demonstrated the existence of the AB effect in a fully two-dimensional system and that the effect of the magnetic field acting on the conductors was to shift the overall (single-slit) diffraction envelope. Although reliable high-field calculations were not possible due to the onset of a numerical instability, results suggest that the oscillations could become damped at sufficiently high fields.

Simulations performed on a larger ring show a great increase in complexity of the wavefunction and it is speculated that quantum chaotic behaviour may be occurring. The results also clearly show a multi-mode structure, even if the impacting wavepacket is entirely within the first transverse sub-band provided the total energy is sufficient to support higher sub-band occupation. It is proposed that it is the formation of these transverse modes which is a dominant factor in reducing the magneto-transmission modulation depth by allowing the side-lobes of the interference pattern to propagate along the output wire irrespective of the enclosed flux.

The overall diffraction envelope in the large-scale simulation was also seen to shift slightly due to the direct action of the magnetic field. Although high-field calculations were again precluded, extrapolation of the available data suggested that one could expect a reduction in AB oscillation amplitude through this mechanism at a field of around 0.5T. This is consistent with experimental results which show the oscillation amplitude starting to decay at fields between 0.5 and 1.0 Tesla.

The transit times of the wavepackets through the ring were also investigated and compared to the analytic results for a single-mode wavepacket. The transit time for wavepackets of energies 1.27 and 10meV propagating through an AB ring was found to be greater than that of an equivalent free wavepacket travelling the same distance in a straight line. This was especially noticeable for the 10meV wavepacket which was slowed appreciably due to the transfer of energy to form the transverse modes across the conductors. Evidence for multiple-transit trajectories was also seen in the transmitted charge versus time plots.

9. Magneto-Resistance studies.

9.1 Introduction

This section describes how the steady-state transmission and reflection coefficients for the ring were calculated and how these were related to the resistance of the ring. This is not only a valuable theoretical study but also an essential step in comparing the theory with experiment.

The results of magneto-transmission calculations for the 400nm diameter ring described in chapter 8 are presented and some general points about the transmission characteristic are noted. Also presented are calculations of the magneto-transmission characteristic for an AB ring fabricated by Ford *et al* [1987]. Calculations are performed for a ring formed from wires 80nm wide using two different wavepacket energies both of which gave a magneto-transmission modulation depth that was larger than those observed in experiments. However, an additional simulation of a similar ring with 200nm wide wires gave a modulation depth that was in better agreement with experimental results.

In order to understand the unexpected features in the transmission characteristics, the simulated data were analysed by fitting to the one-dimensional model (chapter 5) and also by means of a Fourier transform.

Based on the results of this chapter and chapter 8, a technique for increasing the modulation depth of the resistance oscillations is proposed and the results of a wavepacket propagation study for an appropriately modified ring are shown.

Finally some other factors that might affect the magneto-transmission of these rings, such as a longitudinal electric field and random fluctuation potentials, are investigated.

9.2 Determination of transmission and reflection coefficients.

One drawback in using the imaginary-potential technique to model the contacts is that it no longer permits a simple spatial integration of the wavefunction to the left and right of the ring to determine the transmission and reflection coefficients. Instead, the probability current density just before the contacts at each timestep was calculated and then integrated over time to yield the total charges passing into the left and right contacts. The probability current was calculated as:-

$$J = e\hbar (\psi^* \nabla \psi - \psi \nabla \psi^*) / (2mi) - e^2 A \psi \psi^* / m \quad (\text{Eqn 9.1})$$

For the predominantly low-field work performed in this study, the diamagnetic term in the vector potential was negligible. The current just before the contacts was then integrated across the entire y-plane across the conductor to find the transmitted flux for each timestep.

$$\text{Transmitted flux (x)} = \sum_0^{y_{\max}} J(x,y) \Delta y \quad (\text{Eqn 9.2})$$

The reflection coefficient was similarly determined.

9.3 The relation between the transmission coefficient and the conductance.

Having determined the transmission and reflection coefficients, it now remained to convert this into a value for the resistance. Three formulae were considered. The first was the standard tunnelling formula [Tsu and Esaki 1973] which is often used to calculate the current through superlattice type structures and resonant tunnelling devices.

$$J(V) = 2e/(2\pi)^3 \int T(E,V) v(k_z) [f_1 - f_2] dk_x dk_y dk_z \quad (\text{Eqn 9.3})$$

where $f_1 = f(E)$ is the Fermi distribution function at the source contact and $f_2 = f(E + eV)$ that at the drain.

The usual assumption made is that the transverse wavevector is a constant of the motion which allows a polar integration over k_x and k_y to be performed leading to a single integral over the longitudinal component of the energy. From the time-dependent computer simulation it is seen that the transverse component of the wavevector is not conserved in this case and it is not clear how the formula should be changed to take account of this. It is also usually assumed that the transmission coefficient is a function of longitudinal energy only, which again doubtful. The tunnelling formula can be simplified by considering the one-dimensional problem and replacing the 3D density of states prefactor by the 1D form.

$$I(V) = 2e/(2\pi) \int T(E,V) v(k) [f_1 - f_2] dk \quad (\text{Eqn 9.4})$$

Under the assumption of small applied voltages and low temperatures, this form reduces to:-

$$g = e^2/h \int T(E_f) (-df/dE) dE \quad (\text{Eqn 9.5})$$

after a change of variables to integration over energy.

In the zero-temperature limit, the integral can then be performed to yield the one-dimensional conductance formula:-

$$g = e^2/h T(E_f) \quad (\text{Eqn 9.6})$$

which is the same result as that obtained by Economou and Soukoulis [1980] which was derived from linear response theory. This result was later generalised to N quantum channels by Fisher and Lee [1981] and was the second formula considered:-

$$g = e^2/h \text{Trace}(tt^\dagger) \quad (\text{Eqn 9.7})$$

where t is the $N \times N$ transmission matrix expressing the transmission between the N input and N output channels. This formula does however pose a conceptual problem in that the conductance does not tend to infinity as the transmission matrix tends to the unit matrix as one would intuitively expect.

A one-dimensional conductance formula which does yield an infinite conductance as the transmission tends to unity is the Landauer formula:-

$$g = e^2/h\pi \cdot T/(1-T) \quad (\text{Eqn 9.8})$$

where g is the conductance and T is the transmission coefficient evaluated at the Fermi energy. This formula has the advantage of being simple in form, but it does have some provisos. The first is that the simple form is essentially a zero temperature result, although in a later paper [Buttiker *et al* 1985] the result was extended to include finite temperatures. The second, more restrictive, condition is that it is a strictly one-dimensional theory. Despite these restrictions however, the zero-temperature Landauer formula was used in this project as the experimental conditions of very low temperatures (<1K) and small applied voltages (to avoid carrier heating) were such that the zero temperature result could be considered to be a good approximation.

However, this issue is by no means closed and there is considerable debate concerning what constitutes a physically relevant quantum-mechanical expression for the conductance of a disordered conductor [Stone and Szafer 1988].

The conductance of rings with several transverse modes excited leading to several "quantum channels" was also calculated by Buttiker and Landauer [1985] following the same reasoning that lead to the ^{single}channel result. In this study, it was assumed that each of the channels contributed equally to the total transmission and that the Fermi velocity to be associated with each outgoing channel was the same. Proceeding on this assumption, the multichannel Landauer formula reduces to the single channel result and so just the single channel result was used.

The Landauer formula is applicable to the type of situation shown in figure 9.1.

Electrons are assumed to move ballistically from the left reservoir (where electrons are considered to be in thermal equilibrium with the lattice) to the right reservoir and the transmission coefficient together with the 1D density of states is used to find the current passing through the region. The transmission and reflection coefficients were also used to determine the relative shift in the quasi Fermi levels on each side of the region. This can be thought of as finding the voltage generated across the device as a result of an accumulation of charge on the left hand side of the region.

These two steps determined the current flow and the resulting voltage across the device, which in turn determined the resistance. One point to note is that both the steps involved the use of the transmission coefficient, which explains why the Landauer formula is non-linear in the transmission probability.

9.4 The resistance of small rings.

To achieve normal dissipative behaviour in the Landauer model, it was required that the inelastic collision processes occurred in the reservoirs and that the phase of the wavefunction was randomised. Alternatively, if this requirement is lifted, and the wavefunction is coherent in these regions, then it was suggested [Buttiker *et al* 1983] that superconducting effects could occur if the two reservoirs were connected together. Indeed, in such a system the electron would behave in the same way as it would in a periodic potential, and hence should show an energy bandstructure.

Lenstra and Haeringen [1986] challenged Landauer's ideas by showing that "resistive behaviour" could occur in such rings in the absence of inelastic scattering and any phase randomisation as envisaged by Landauer. Lenstra and Haeringen's analysis was based on the fact that the energy eigenvalue spectrum for a ring enclosing a magnetic flux consists of a series of intersecting bands, and in the presence of a slight amount of elastic scattering, electrons could change bands at the intersections. This provided a mechanism by which electrons could follow many different trajectories in phase-space to arrive at the same point. It was then stressed that if there were a large enough number of intersecting bands, the arrival phases of the electrons would be indistinguishable from a random set of angles distributed between 0 and 2π . A numerical calculation showed that there was a near-linear relationship between the current and the field. However, a linear relationship between current and field alone is not sufficient to prove the existence of dissipative resistance. For instance, one could imagine a purely ballistic device in which electrons are emitted from a source at a rate such that the local electron density remains a constant (like a vacuum tube). In this device, the current is proportional to the applied voltage across the system, but no dissipation occurs. In the case of a vacuum tube, dissipation only occurs when the electrons collide inelastically with the anode (neglecting radiative losses). To obtain Ohms law behaviour as it is usually understood, requires the electron gas to be close to thermal equilibrium with the lattice. In devices where transport is ballistic, the *electron gas* becomes hot with respect to the lattice. That is,

the externally provided energy is stored in the form of kinetic energy of the carriers, but is not dissipated in the form of heat in the lattice. To obtain normal resistive behaviour, there seems to be no avoiding the eventual need for electrons to experience inelastic collisions with the lattice; a view which was also taken by Buttiker [1986].

The argument of Lenstra and Haeringen does, however, question whether merely ascribing a phase randomisation to the wavefunction ensures dissipative behaviour. Their model was completely deterministic, but nevertheless produced a combination of phases in the wavefunction which were indistinguishable from a true randomising process such as an inelastic collision. Thus, it would seem that an inelastic collision causing energy dissipation causes a phase randomisation in the wavefunction, but the identification of a random (or seemingly random) phase in the wavefunction does not necessarily mean that an inelastic collision has taken place.

In the analysis of Landauer, inelastic processes are invoked in the contact regions and provide the necessary energy dissipation. Landauer then goes on to deal with the consequences of this, namely phase randomisation in the contacts.

9.5 The magneto-transmission characteristic of a 400nm diameter ring.

Figure 9.2 shows the transmission coefficient of a 10meV wavepacket through an AB ring of 400nm outside diameter and a wire width of 66.4nm. The wavepacket, which had the same parameters as the one used in Chapter 8, was propagated for 50ps by which time the interaction with the ring was judged to be complete. This calculation was repeated for different magnetic flux values to obtain the magneto-resistance characteristic. The first half of the graph consists of 13 calculated points and for the second half and all subsequent graphs, 10 points per h/e cycle was considered to be sufficient. A smooth line through these points was then drawn using a cubic-spline interpolation procedure [Press *et al* 1987], which ensures continuity in the first and second derivatives from one point to the next.

There are two minima in the transmission characteristic of 0.272 and 0.236 which occur at the $1/2(h/e)$ and $3/2(h/e)$ positions respectively. Although these minima are near the positions one would expect from simple theory, they are not exactly at the

$1/2(h/e)$ and $3/2(h/e)$ positions. This could be due to the magnetic field acting on the wavepacket, causing a shift in the overall envelope. Alternatively, it could be because the assumed average radius of the ring experienced by the wavepacket is not simply the average of the outer and inner ring radii, which would mean that the true period would be different to the value assumed for the calculation. Of these theories, the magnetic-shift explanation seems the more plausible as there is good experimental evidence [Ford *et al* 1987] to suggest that the h/e flux period is associated with the average of the outer and inner radii of the ring. In the context of this study there was no reason to suspect the original calculated value for the flux period.

Viewed as a whole, there is a downward trend in the magneto-transmission. This can be explained by referring to figure(8.58) which shows a plot of the probability density for a flux value of h/e . If no magnetic force was acting on the wavefunction, then this plot would be indistinguishable from a plot at the same time slot for a flux value of zero. It is seen that whilst the plot for flux value h/e is similar to that for zero flux, the central maximum has been shifted clockwise by a slight amount. The lower fringe is more effectively obscured by the output wire, reducing the contribution of that fringe to the total transmission and at the same time the transmission of the upper fringe has been virtually unaffected by the shift. Since the second-order fringe in the upper arm has not yet reached a position where it can contribute towards the transmission, the overall transmission decreases. It is thought that the transmission will start to show an upward trend when the second-order fringe starts to contribute and the lower first-order fringe becomes completely cut-off.

The initial rise in transmission is very different to the cosine-like curve predicted from simple theory. It does however bear some resemblance to the transmission characteristic predicted by the 1D transmission model. Thus, the initial rise in the transmission seen in the two-dimensional model is likely to be due to reflection processes taking place at the input and output junction of the device. This question will be addressed in the following section.

The origin of the small perturbations in the curve is rather more difficult to ascertain. One explanation would be that they are the result of a superposition of a cosine-like and a 1D-like transmission characteristic (showing resonances). This hypothesis is later proposed as an explanation for the shape of the minima of the magneto-transmission of other rings.

Figure 9.1 Use of the Landauer formula to relate the transmission coefficient of a one-dimensional potential to the conductance.

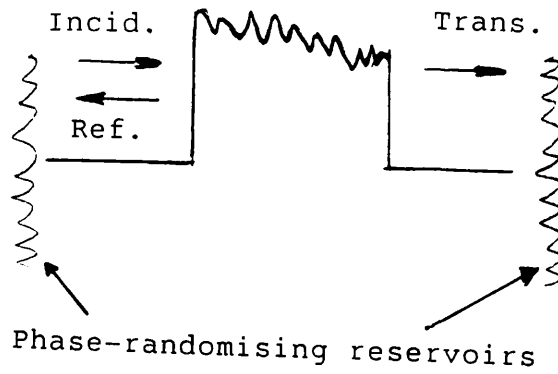
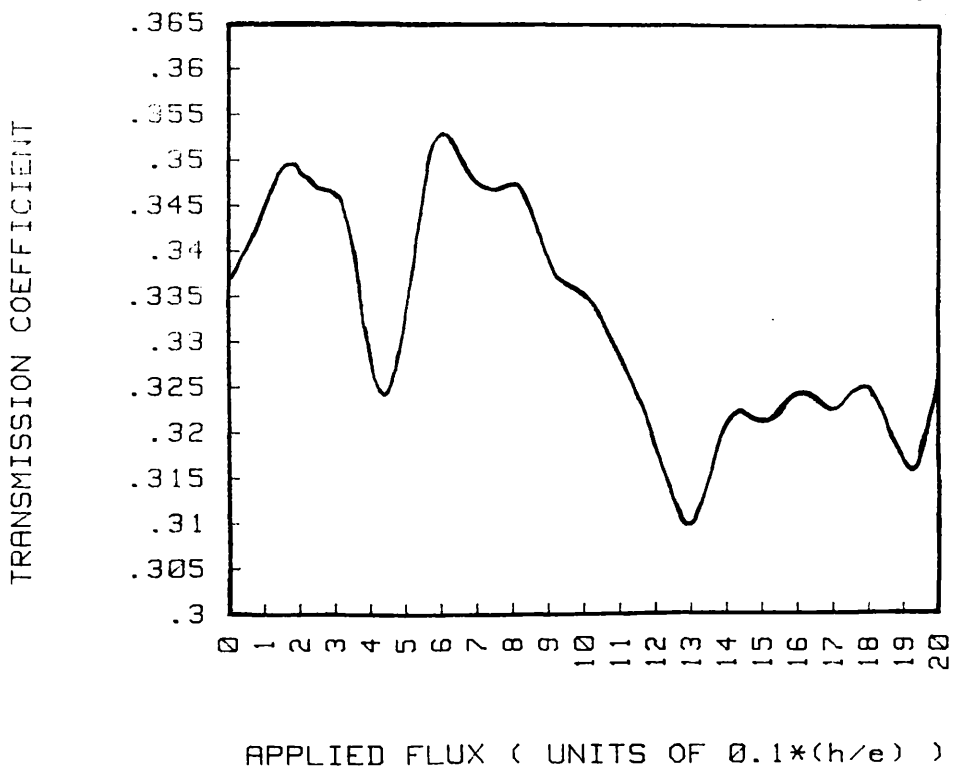


Figure 9.2

MAGNETO-TRANSMISSION OF A 4000 ANGS. DIAMETER A-B RING
Conductor width=66.4nm



9.6 Comparison with one-dimensional analysis and experimental results.

9.6.1 General.

At the time this work was in progress rings fabricated using the shallow-etch technique had not been perfected and so it was therefore decided to model rings previously fabricated and compare with published experimental results.

The devices fabricated by Ford *et al* [1987] who used the gated HEMT technique were chosen in this instance. Much experimental data were available for these rings and the amplitudes of the AB oscillations were initially around 5% of the background resistance. Unfortunately, these structures were about 3 times larger than those initially anticipated using the shallow etch technique.

Practical considerations on memory space and CPU time on the computer restricted the number of grid points to be about the same as that used in the simulation of the 400nm ring. This meant that the size of each mesh cell had to be increased accordingly, which sacrificed accuracy. Nevertheless, the results were still expected to be about 5% accurate or better. The dimensions of the Cambridge ring were taken to be:-

Outside radius of the ring = 0.69 micron

width of wire forming the ring = 80nm

The average radius of 0.65 micron, deduced from the lithographic dimensions, was consistent with the position of the h/e peak in the Fourier transform of the magneto-resistance data and the wire width of 80nm was an estimate based on the width of the h/e peak. From measurements of the electron concentrations in the 2DEG when configured as an AB ring, the Fermi energy was estimated to be between 9 and 13meV.

9.6.2 Analysis of magneto-transmission results in terms of hybrid 1D and double-slit models.

Figures 9.3 and 9.4 show the magneto-transmission characteristics for wavepackets of 9 and 13meV average energy respectively incident on the ring described above. There are significant differences between the two curves. The magneto-transmission curve for the 9meV wavepacket has an average transmission of 0.309 compared to 0.111 for the 13meV wavepacket. It was originally thought that increasing the average wavepacket energy would cause an increase in the average transmission coefficient and a decrease in the modulation depth of the oscillations due to the larger number of modes which would be sustainable across the wire. The transverse sub-band energies for an 80nm infinite potential well suggest that a significant occupation of the fourth sub-band can be expected for the 13meV wavepacket but only a very small occupation in the case of the 9meV wavepacket, which would seem to support this view.

One possible explanation of the observed behaviour in the simulated characteristics is that the one-dimensional character of the ring is causing the magneto-transmission curve to exhibit energy dependences such as those found in chapter 5. The problem with this interpretation is that a low average transmission implies a low coupling coefficient for the ring, which in turn leads one to expect sharper structure in the transmission characteristic than is actually seen. However, it is possible that any sharp resonances present in the 13meV curve may have been missed, as the magneto-transmission curves were constructed at a resolution of only 10 points per h/e cycle. In order to investigate the extent to which the simulated ring exhibited one-dimensional behaviour, the data were fitted to a 1D transmission characteristic.

One of the problems encountered in the one-dimensional theory was that there was no method of determining the coupling parameter for the S-matrices so the coupling parameter was left as an adjustable parameter in the theory. The two-dimensional simulation however, provided a method of approximately determining the coupling parameter for a ring of a more realistic geometry.

The fitting procedure began with finding a value for the wavevector-length product which was calculated from the average wavepacket energy and the length of one of the arms of the ring from input to output using the average radius. The coupling parameter was then found by matching the transmission of the 1D model to the 2D magneto-transmission curve at zero flux.

Figure 9.5 shows a plot of the transmission versus the enclosed flux using the 1D model with model parameters taken from the 2D magneto-transmission curve for the 9meV wavepacket. The parameters are $kl=0.88*2\pi$ and $\varepsilon=0.305$ for the wavevector-length product and coupling parameter respectively.

The 1D transmission characteristic using these model parameters is very different to that obtained from the simulations (such as that shown in figure 9.3) showing that the 1D model is not even a good approximation to the 2D situation. However, some of the small perturbations in the simulated curves may be due to one-dimensional effects. For instance, the 9meV curve has a much sharper dip in transmission around the $1/2(h/e)$ positions than the cosine curve expected from simple theory.

A qualitative explanation of the shape of the curve for the 9meV wavepacket (and the 10meV wavepacket incident on the 400nm ring) can be gained by considering the modelled ring to be intermediate between an open double-slit type geometry and a purely one-dimensional geometry. The open-geometry case is that in which the wires forming the ring are considered to be infinitely wide, but in which the transmission is still measured at a single point at the output, so that this case represents a different type of multiply-connected 1D geometry. In this case the simple cosine form, discussed in chapter(3), can be used for the transmission.

$$T(\phi)=1/2(1+ \text{Cos}(2\pi \phi/\phi_0)) \quad (\text{Eqn } 9.9)$$

whereas for the 1D geometry, the model developed in chapter 5 must be used.

To develop a hybrid model for the simulated ring, it was assumed that the transmission could be modelled by a linear combination of the double-slit and 1D transmission forms such that the overall result is normalised:-

$$T(\phi)= \alpha T_{1D}(\phi) + (1- \alpha) T_{\text{open}}(\phi) \quad (\text{Eqn } 9.10)$$

where α is a parameter representing the degree of one-dimensional behaviour.

Figure 9.6 shows the result of such a fitting for a 9meV wavepacket incident on the Cambridge ring of 1380nm diameter and 80nm wide wires. The best fit for the parameter α , for the calculated kl product, was found to be about 0.2. It can be seen that the general shape of the dip in transmission can be reproduced using this model,

Figure 9.3 Conductor width=80.0nm Wavepacket energy=9.0meV

MAGNETO-TRANSMISSION OF A 690 nm OUTSIDE DIAMETER A-B RING

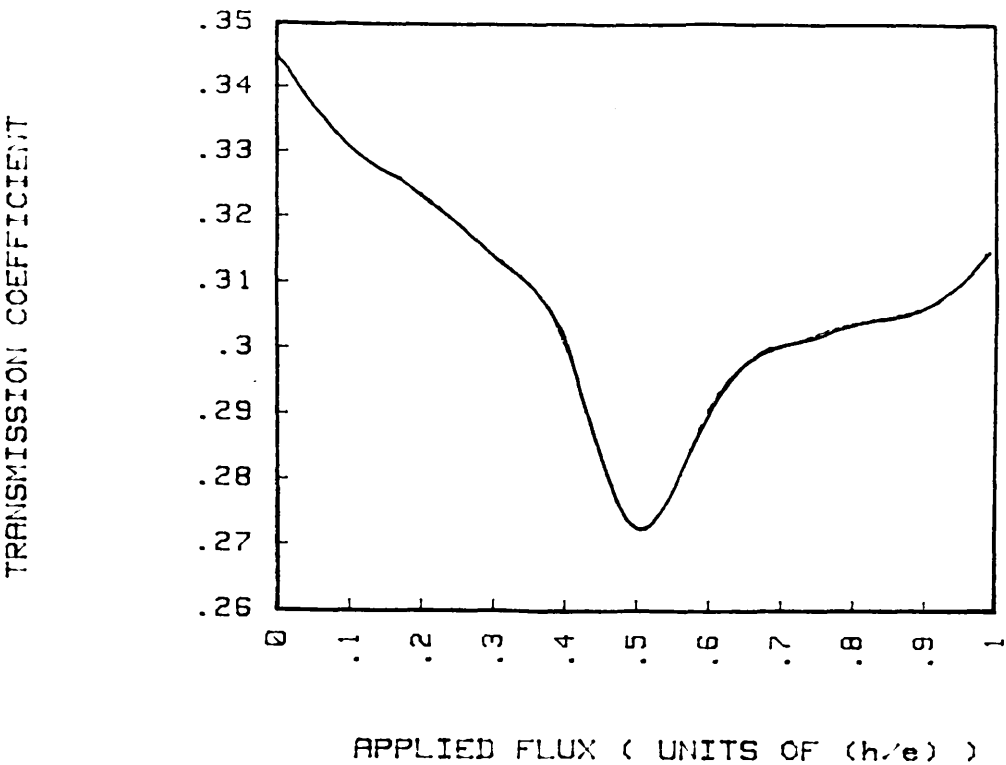


Figure 9.4 Conductor width=80.0nm Wavepacket energy=13.0meV

MAGNETO-TRANSMISSION OF A 1380nm DIAMETER A-B RING

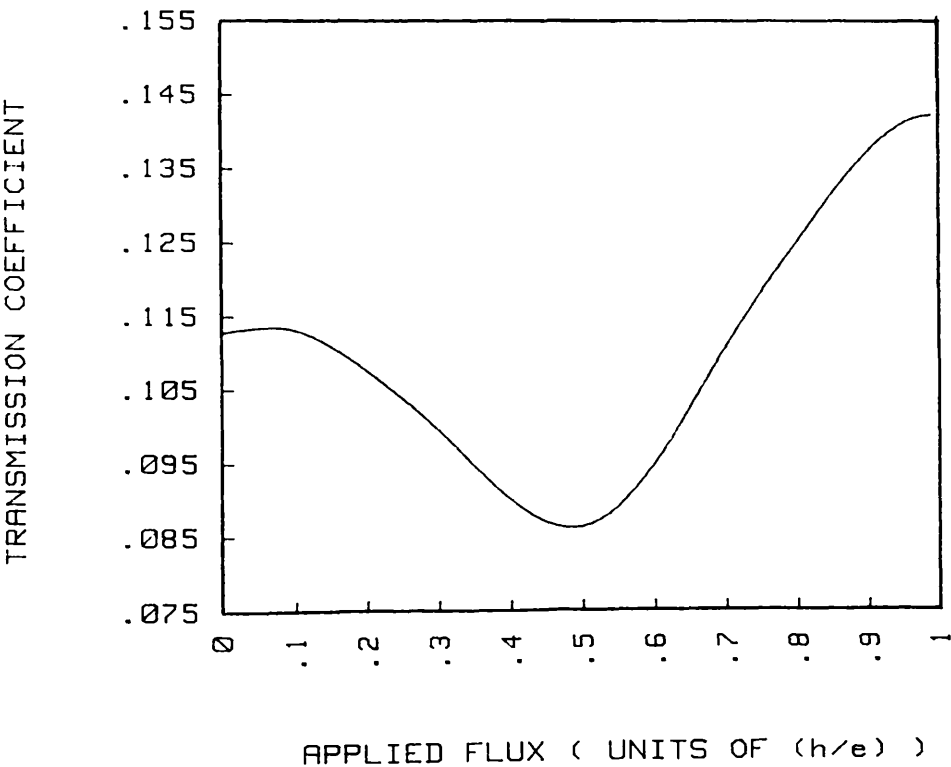


Figure 9.5 1D transmission model using data taken from
the 2D magneto-transmission curve of the 400nm ring

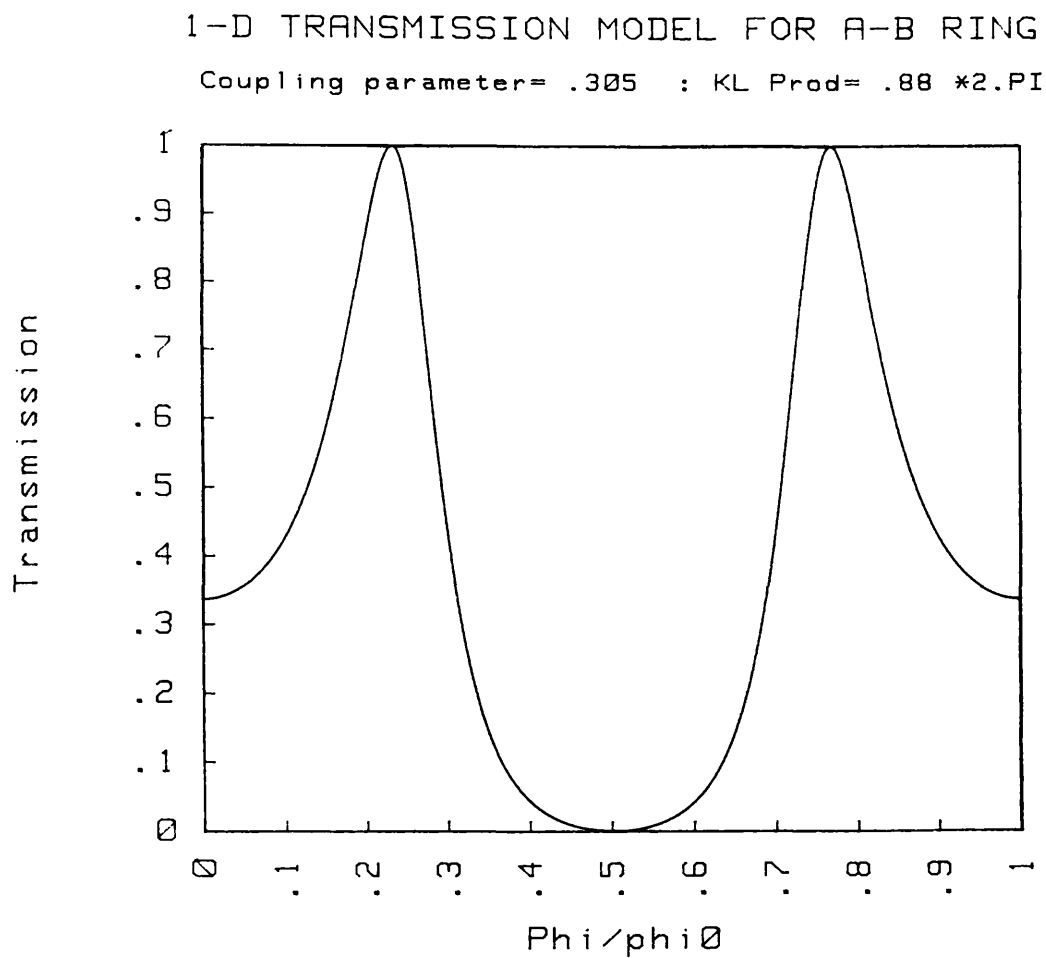


Figure 9.6

Hybrid 1D transmission model to fit the 2D magneto-transmission data from the Cambridge ring with a 9.0meV incident wavepacket energy.

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .305 : KL Prod= .88 *2.PI

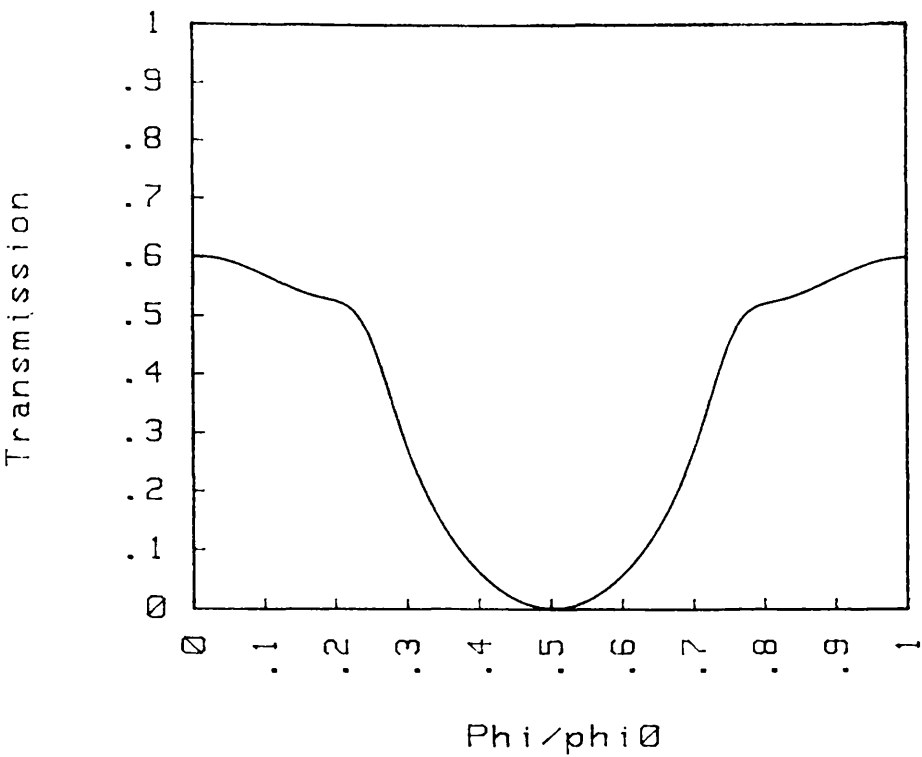


Figure 9.7

Improved hybrid 1D transmission model for the Cambridge ring using a one-dimensionality parameter of 0.2

1-D TRANSMISSION MODEL FOR A-B RING

Coupling parameter= .3 : KL Prod= .8 *2.PI

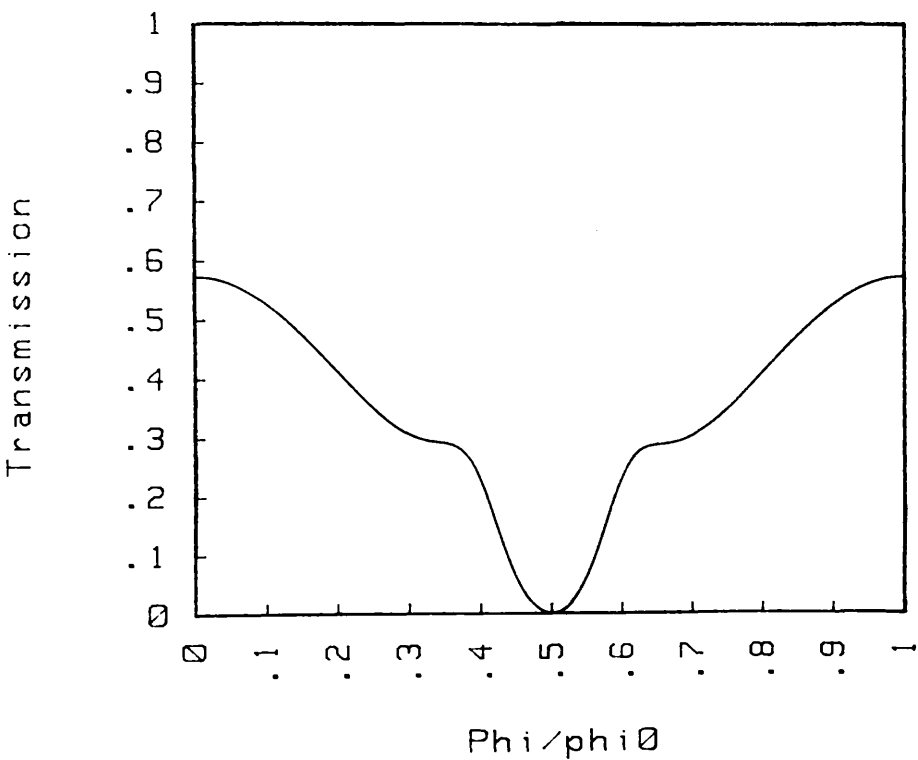


Figure 9.8 Hybrid 1D transmission model for the Cambridge ring using a one-dimensionality parameter of 0.3

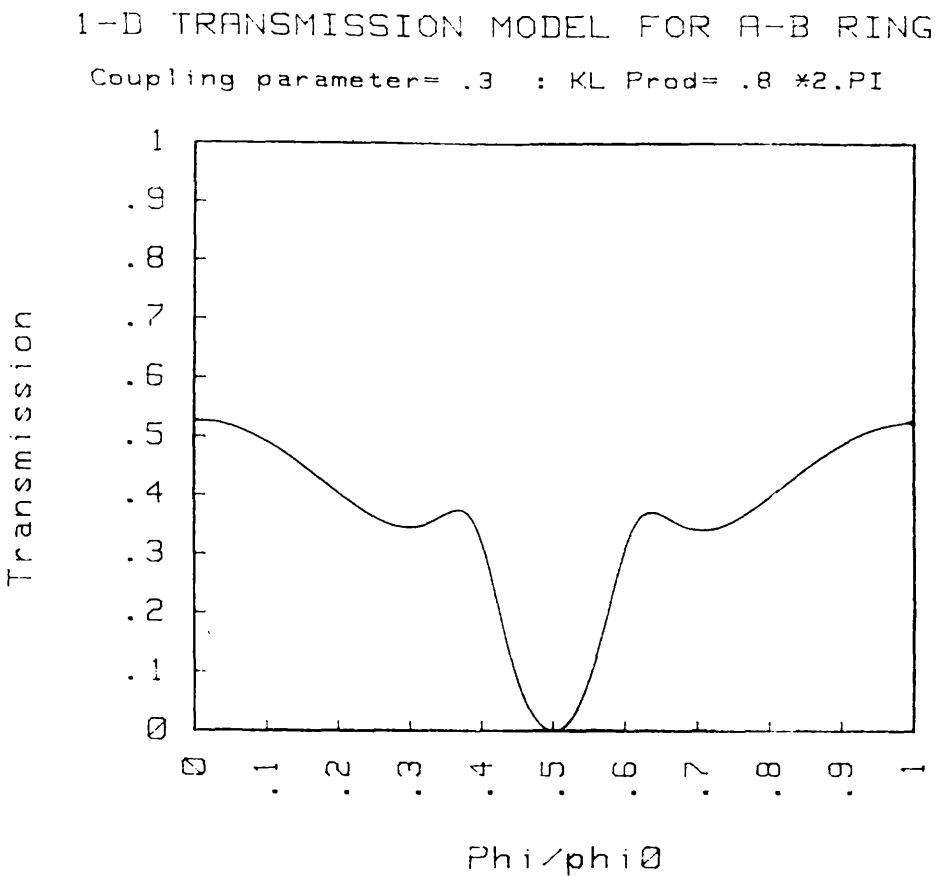
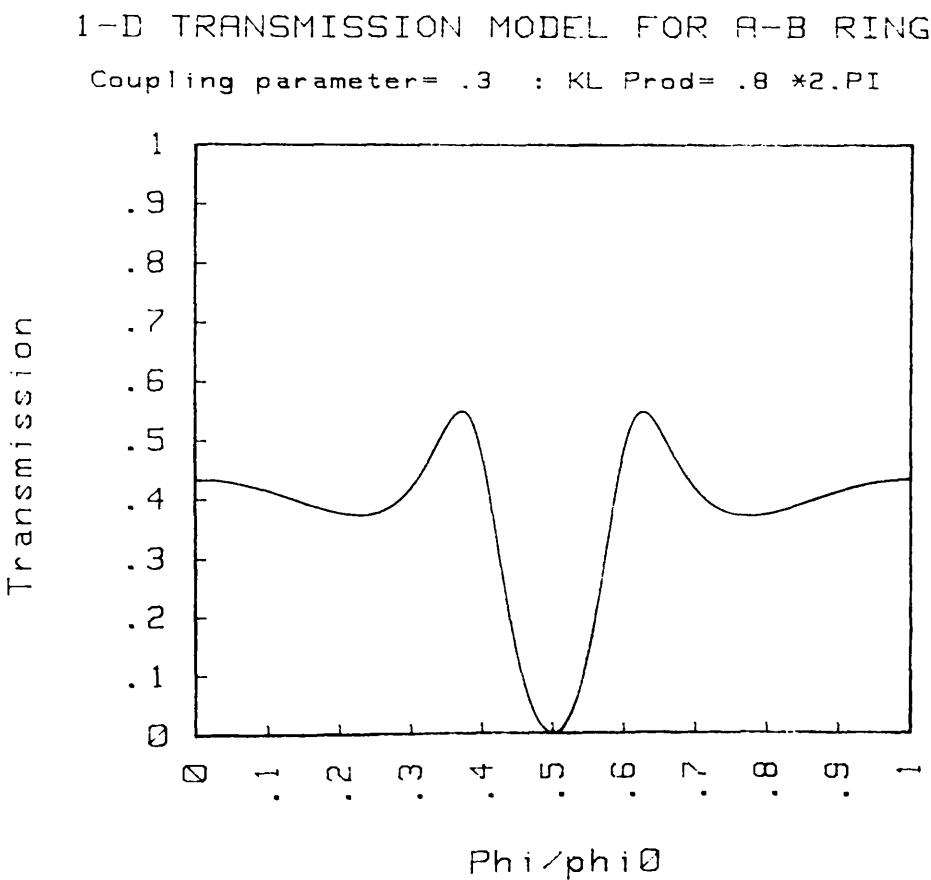


Figure 9.9 Hybrid 1D transmission model for the Cambridge ring using a one-dimensionality parameter of 0.5



Figures 9.10 and 9.11: The effect of changing the coupling parameter in the hybrid 1D transmission model. The one-dimensionality parameter is 0.2

Figure 9.10

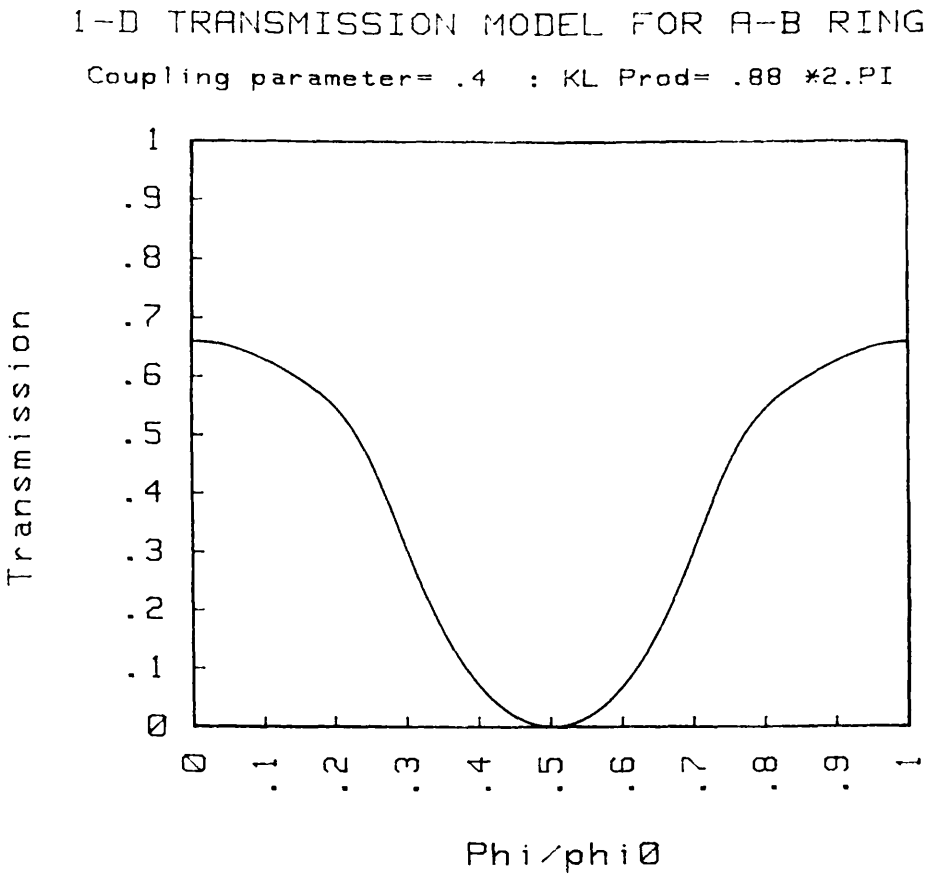
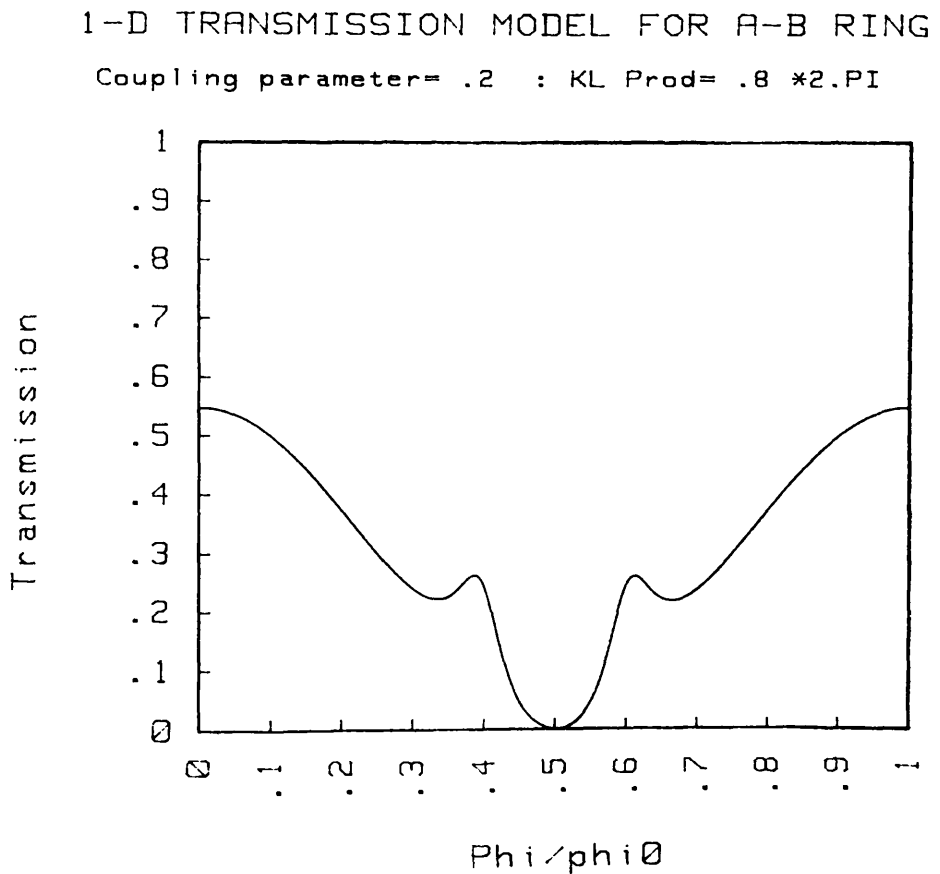


Figure 9.11



although the transmission always reaches zero when the enclosed flux is $1/2(h/e)$ because both components of the hybrid model are one-dimensional. The non-zero transmission in the simulated curve can therefore be directly attributed to the two-dimensionality of the ring. The hybrid model also shows that the simulated curve can be better understood in terms of the 1D-type resonances enhancing the transmission either side of the $1/2(h/e)$ minimum, rather than the dip itself being thought of as sharper than expected.

A variation of the fitting parameters showed that the shape of the hybrid curve was relatively insensitive to the value of the coupling parameter and the one-dimensionality parameter α , but was strongly dependent on the wavevector-length product kl .

It was found that the calculated kl value of $0.88 \cdot 2\pi$ reproduced the shape of the simulated curve reasonably well, but a slightly lower value of 0.8 and a coupling parameter of 0.3 was considered to give a better fit. Figures 9.7 to 9.9 show the hybrid curve for these parameters and also the effect of varying the parameter α . The effect of increasing α was to sharpen the dip in transmission at the $1/2(h/e)$ flux value and to enhance the peaks either side of the dip. Figures 9.10 and 9.11 show the effect of changing the coupling parameter to 0.2 and 0.4 respectively whilst keeping $\alpha=0.2$ and $kl=0.8 \cdot 2\pi$.

9.6.3 Harmonic analysis of the simulated magneto-transmission characteristics.

A second technique used for assessing the structure of the magneto-transmission curves was to Fourier transform the data. An electron trajectory passing either side of the ring once only and interfering at the output will give rise to a fundamental oscillation frequency in the transmission of period h/e whilst those which continue around the ring and interfere at the input contribute a frequency component of period $h/2e$ because they have encircled twice as much flux as the single h/e trajectories. Depending upon how much reflection takes place at the input and output junctions there will also be higher harmonics present in the Fourier spectrum due to multiple-transit trajectories. A Fourier transform can therefore be expected to provide information on

the importance of the various types of trajectories. Figure 9.12 shows a cosine Fourier transform of the magneto-transmission data for one h/e cycle for the Cambridge ring using a 9meV incident wavepacket energy after the average (background) transmission had been subtracted.

It is seen that the dominant frequency component is the 2nd harmonic, which is due to the double-passage trajectory and can be attributed to the sharp dip in the magneto-transmission data. Thus, in terms of the one-dimensional analysis, it can be said that the ring is quite poorly coupled to the external leads.

Figure 9.13 shows a Fourier transform of one h/e cycle of the magneto-transmission data for the 400nm ring using a 10meV incident wavepacket. In this spectrum the h/e fundamental is the dominant frequency component, but there is still a large contribution from the 2nd harmonic and also a surprisingly large 5th harmonic, which could be subject to some exaggeration due to the spline fitting technique used.

Figures 9.14 and 9.15 show the Fourier transform for two h/e magneto-transmission cycles for the 1380nm Cambridge ring and the 400nm ring respectively. The transform of the data for the 400nm ring still shows the first and second harmonic frequencies to be large, but their amplitudes relative to the higher frequency components has been diminished. This is probably because of the proportionally larger high-frequency components needed to describe the overall negative slope in the transmission curve.

Due to the fact that only one or two h/e cycles of data were transformed, it would be inappropriate to infer any wire-width dependences from the width of the Fourier peaks as the resolution in the Fourier domain was judged insufficient. However, the general form of the Fourier spectrum may provide some additional information. The transforms of all the magneto-transmission curves show a decrease in the Fourier amplitudes with increasing frequency, the envelopes of which can be seen to resemble an exponential decay. This behaviour is also seen in experiment and is thought to be due to the increased probability that electrons giving rise to the high-frequency components will suffer a phase-breaking collision. However in this simulation no phase-breaking collisions have been included, so this mechanism cannot be responsible for the observed characteristics of the simulated ring. Nevertheless, the approximate exponential decay of the Fourier components can be understood in this

ballistic model by making the following hypothesis. Suppose that the portion of the initial wavepacket which has passed into the ring then oscillates back and forth from the input junction to the output junction of the ring. Suppose further that if upon each rebound there is a constant probability of some fraction of the trapped wavefunction escaping from the output or input lead. One can then see that on each rebound the absolute contribution for multiple-transit trajectories diminishes approximately exponentially because the trapped wavefunction amplitude available inside the ring for transmission has been reduced by previous rebounds.

9.6.4 Comparison of simulated results with experiment.

Figure 9.16 shows the first few magneto-resistance cycles for the Cambridge ring [Ford 1988] obtained experimentally at a temperature of 100mK. It seen that the behaviour of the real ring is quite complex and apparently contains more structure than the simulated characteristics. This is perhaps not too surprising in view of the fact that several simplifying assumptions have been made such as the neglect of scattering on impurity fluctuation potentials and of self-consistent potentials. However even if all the relevant processes were included in the model, the details of the resistance fluctuations would depend on the microscopic arrangement of scattering sites in the device and would therefore change from sample to sample. The development of such a model would be possible in principle but ultimately not very useful. A more useful approach is to consider the "average properties" of these devices and to compare the resistance characteristics on a statistical basis. It was therefore decided to compare the average amplitudes and periods of the fluctuations without trying to reproduce the precise details of the experimental curves.

In simulating the propagation of a wavepacket through the ring, the effects due to quantum interference processes are taken into account but any incoherent processes will not be catered for. Also, parallel conduction in the highly-doped AlGaAs layer has been neglected, so the simulated results must be viewed as a quantum correction to the resistance of the ring.

The oscillation period of the fluctuations in the simulated data (3.12mT) were in good agreement with the experimental data (2.84 +/-0.1mT), but the oscillation amplitude found by simulation was substantially larger than the average amplitudes

found experimentally. The average transmission of the simulated ring was about 0.29 taken over two cycles and the change in magneto-transmission per cycle was approximately 0.04. From these data, the single-channel Landauer formula gave a percentage change in the resistance of the sample ($\Delta R/R = \Delta T / (T(1-T))$) of about 20-30%. The experimental value quoted for a large number of oscillations was much lower than this; being about 5% of the background resistance. The simulated data using the 13meV wavepacket gave an even larger magneto-resistance fluctuation amplitude of about 50%. The increase in magneto-resistance fluctuation for the 13meV wavepacket was aided by the fact that the transmission had fallen substantially, which increased the sensitivity of the ring (via the Landauer formula) to changes in transmission. The reason why the transmission fell so markedly however is not clear. One possibility is that a resonant energy had been reached where a significant amount of reflection occurred at the output junction, although from the data available this was difficult to substantiate.

Figures 9.17 and 9.18 show the magneto-resistance characteristics for the 9meV and 13meV wavepackets respectively. At zero magnetic field, the resistances of the rings are 31.6 and 103k Ω respectively, which are both higher than the experimental value of 10.5k Ω which means that the discrepancy cannot be accounted for by assuming an additional contact or bulk resistance to be in series with the ring.

In further experimental investigations on these ring structures [Ford *et al* 1988] doubt was cast on the technique of using the width of the h/e Fourier peak to measure the wire width. It was suggested that because of the specular nature of the collision of the electrons with the side-walls of the ring, the full range of area-enclosing trajectories would not be explored giving rise to an underestimate of the wire width. Further evidence for this inaccuracy came when it was found that the wire widths inferred from magnetic depopulation of the 1D sub-bands, a technique which had been successful in 1D wires [Berggren *et al* 1986], consistently gave larger widths than those inferred from the Fourier technique. The wire width inferred from magnetic depopulation studies of a similar ring to the one modelled gave a value of about 150-200nm and was thought to be a more reliable figure than that obtained by Fourier analysis of the magneto-resistance.

The wider wires were expected to have two consequences for the magneto-transmission characteristic.

1) To decrease in the percentage resistance oscillation due to a greater number of interference fringes across the wire.

2) To decrease in the zero-field resistance arising from the greater width at the input and output junctions which would increase the transmission.

Figures 9.19 and 9.20 show the results of a simulation for a ring of 1380nm diameter, wire width 200nm and a revised wavepacket energy of 7.0meV [Ford *et al* 1988] to reflect the slightly lower sheet electron concentration. The percentage resistance fluctuation calculated for this ring using the single-channel Landauer formula is about 12%, which verifies the expectation that the resistance modulation amplitude should decrease with wider wires and is in better agreement with the experimentally quoted value of 7%.

In an effort to increase the magneto-resistance oscillation amplitude, a ring with a modified geometry was fabricated [Ford *et al* 1989 and Ford 1988]. This ring, whose geometry was described in detail in chapter 4 had a narrower exit and narrower conductors than the first ring which was expected to reduce the number of modes present. The maximum value of the magneto-resistance oscillation amplitude was about 20% of the background resistance.

Although the ring diameter was slightly different for this ring, magnetic depopulation studies showed that the conductor width was about 80-100nm which was compatible with the width of 80nm considered in the first set of calculations. The experimental value of 18% of the background value for the magneto-resistance oscillation amplitude is now seen to be in better agreement with the initial simulated value of between 20 and 30%.

Figure 9.12 Fourier transform of one magneto-transmission cycle from the Cambridge ring using a wavepacket energy of 9.0meV.

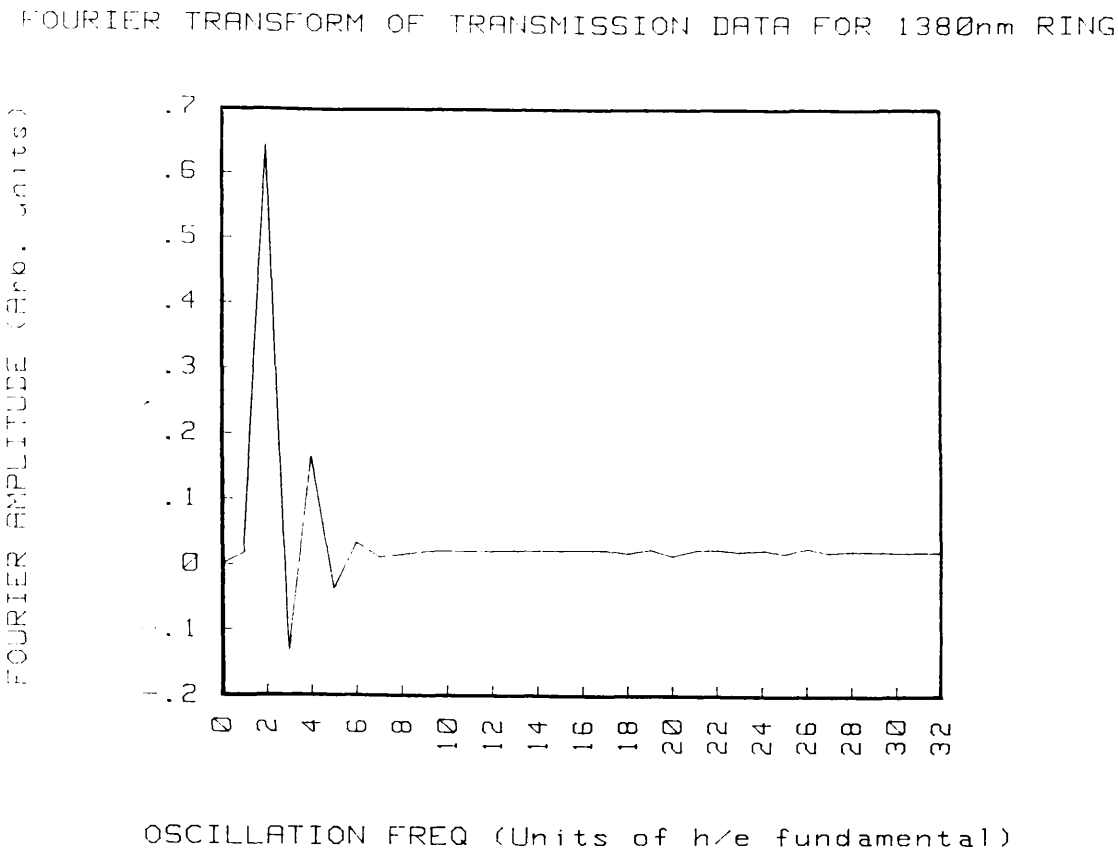


Figure 9.13 Fourier transform of one magneto-transmission cycle from the 400nm diameter ring using a wavepacket energy of 10.0meV.

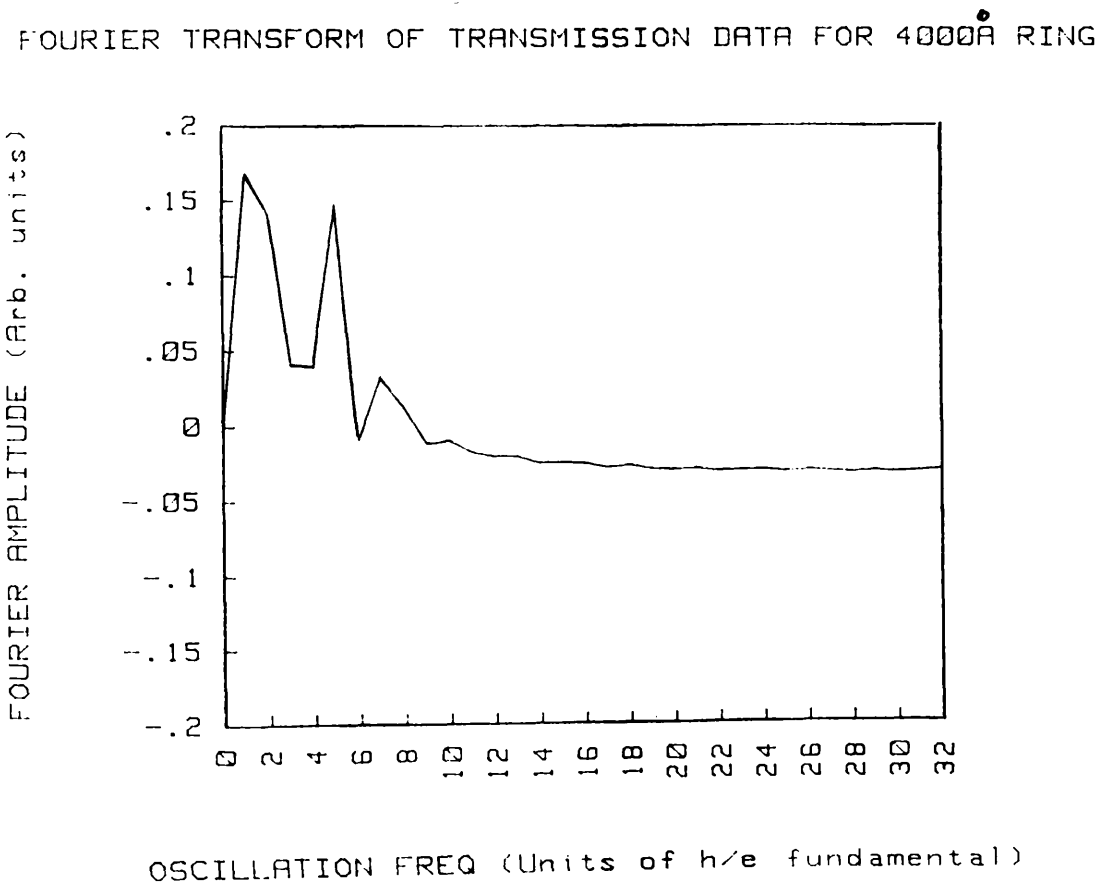
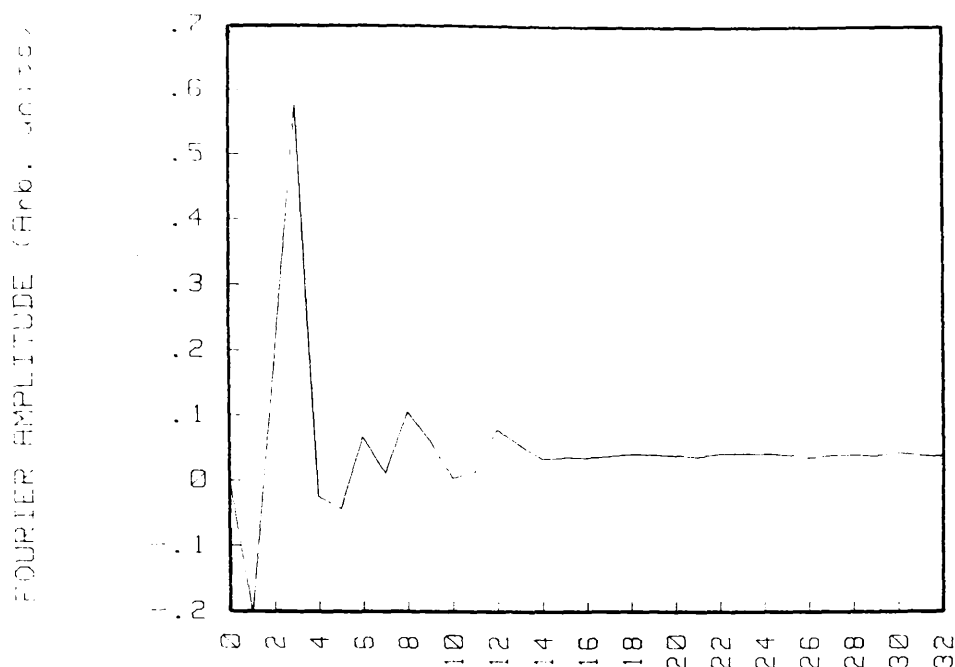


Figure 9.14

Fourier transform of two magneto-transmission cycles from the Cambridge ring using a wavepacket energy of 9.0meV.

FOURIER TRANSFORM OF TRANSMISSION DATA FOR 1380nm RING

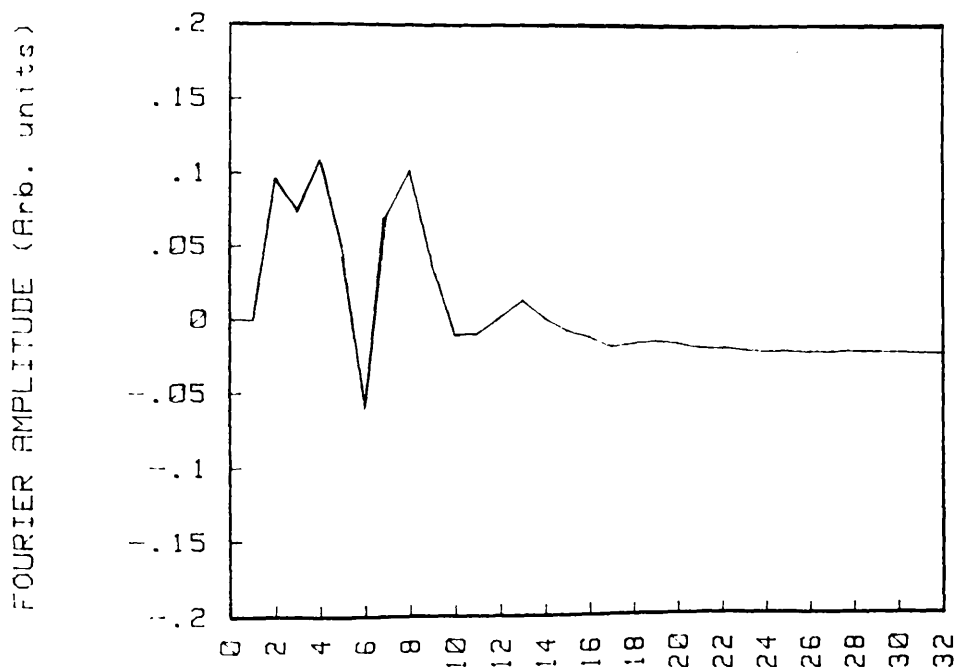


OSCILLATION FREQ (Units of 0.5*h/e fundamental)

Figure 9.15

Fourier transform of two magneto-transmission cycles from the 400nm diameter ring using a wavepacket energy of 10.0meV.

FOURIER TRANSFORM OF TRANSMISSION DATA FOR 4000A RING



OSCILLATION FREQ (Units of 0.5*h/e fundamental)

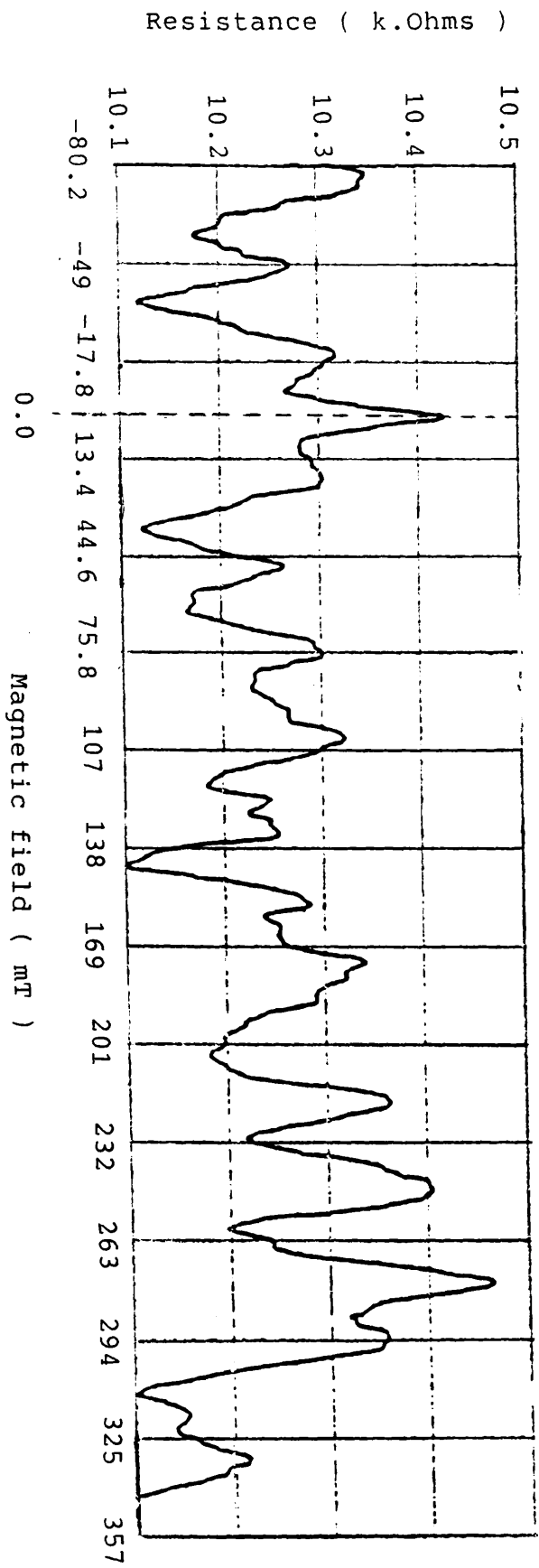


Figure 9.16 Experimental magneto-resistance data
from the Cambridge 1380nm diameter ring.

Figure 9.17 Magneto-resistance of the Cambridge ring using an incident wavepacket energy of 9.0meV.

MAGNETO-RESISTANCE OF A 1380nm OUTSIDE DIAMETER A-B RING

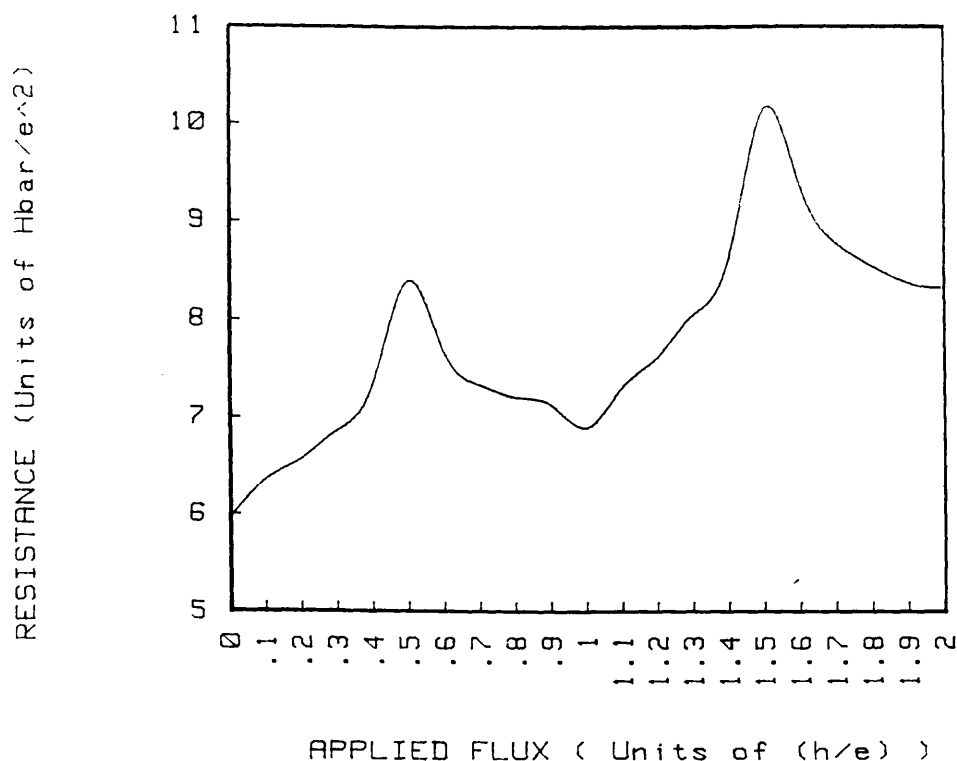


Figure 9.18 Magneto-resistance of the Cambridge ring using an incident wavepacket energy of 13.0meV.

MAGNETO-RESISTANCE OF A 1380nm DIAMETER A-B RING

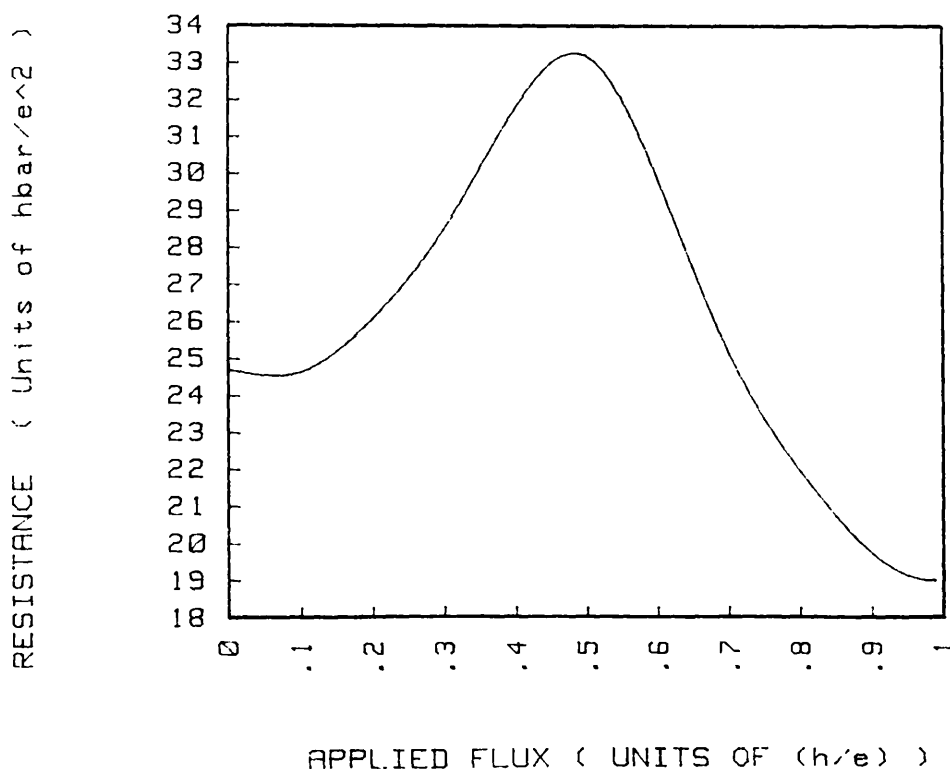


Figure 9.19 Magneto-transmission of the Cambridge ring
 formed from 200nm-wide conductors and taking
 a wavepacket energy of 7.0meV.

MAGNETO-TRANSMISSION OF A 1380nm OUTSIDE DIAMETER A-B RING

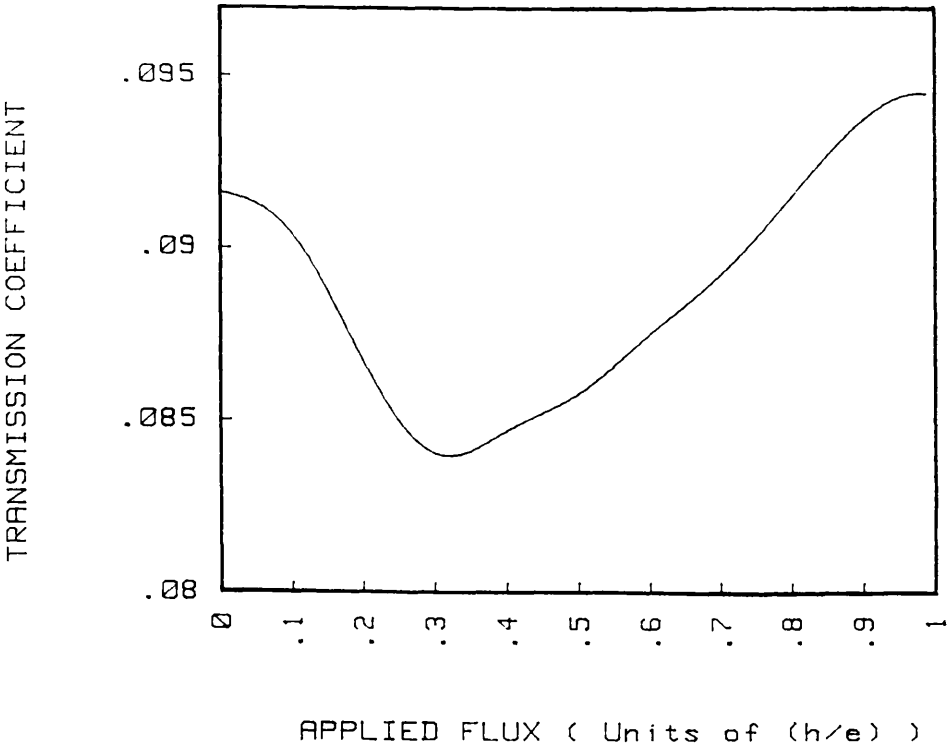
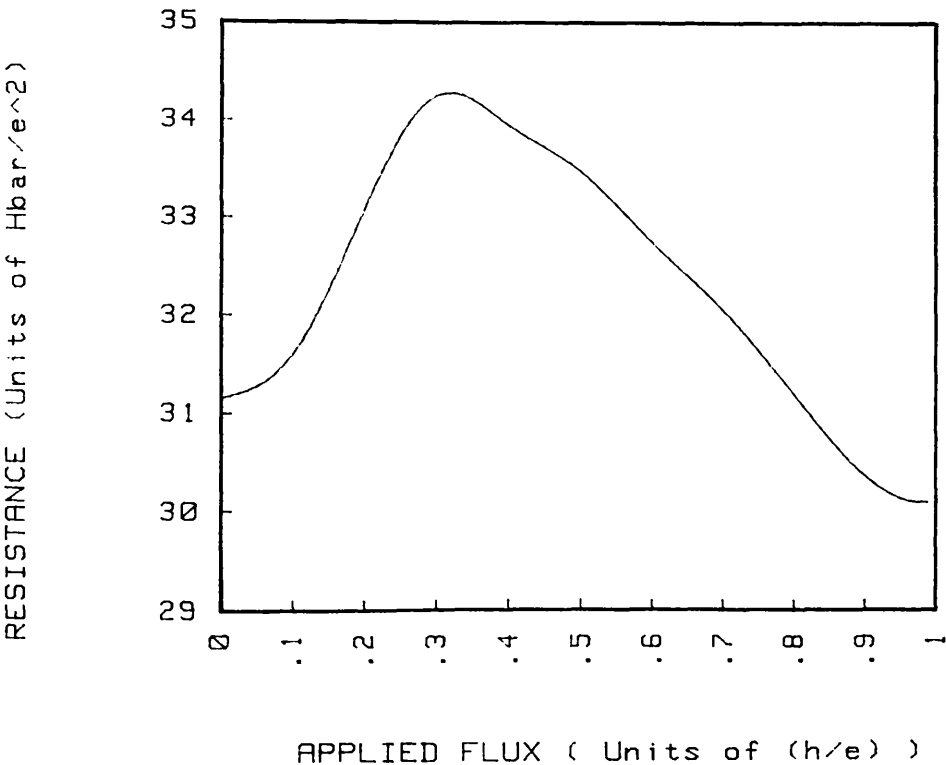


Figure 9.20 Magneto-resistance of the Cambridge ring
 formed from 200nm-wide conductors and taking
 a wavepacket energy of 7.0meV.

MAGNETO-RESISTANCE OF A 1380nm OUTSIDE DIAMETER A-B RING



9.7 A method of increasing the magneto-transmission modulation depth.

Having identified the transmission of the side lobes in the interference pattern as being responsible for the loss of modulation depth in the magneto-transmission characteristics, a method for eliminating this effect was investigated.

The objective was to use the computer model to analyse the behaviour of a wavepacket incident on a ring whose exit was constricted so that only one mode could propagate along the output wire. The initial wavepacket was chosen so that a number of transverse modes were able to propagate in the rest of the ring, as this would invariably be the experimental situation.

The propagation of a 10meV wavepacket through a 400nm diameter ring which was the subject of study in chapter 8 was again simulated but this time with a constriction at the exit of the ring. The constriction was arranged so that the first 80nm of the output wire was only 22nm wide compared to the unconstricted width of 66.4nm, thus only allowing the central peak to be co-incident with the output gap.

Figures 9.21 to 9.40 show a time-sequence of contour plots of the probability density for the modified ring with no magnetic flux applied. The contour plots show that the initial part of the wavepacket propagation to be similar to that of the unmodified ring, the major difference occurring when the partial wavepackets start to interfere at the exit of the ring. It is seen that only a single mode (due to the central interference fringe) propagates through the constriction. There is also an accumulation in the probability density in the constricted region which is the result of a decrease in the forward component of the energy in that region.

Figures 9.41 to 9.46 show selected frames from a time-sequence of contour plots for the 400nm ring but this time enclosing half an h/e quantum of magnetic flux. These contour plots show that with the central peak now absent, a negligible proportion of the wavepacket propagates through the constriction.

The asymptotic transmission coefficient for this ring with no applied magnetic field was found to be 0.1456 and when half a flux quantum was applied, this value fell to 1.658×10^{-2} . The ratio of the "on" to "off" transmission with the constriction is about 8.8 compared to an on/off ratio of about 1.04 for the unconstricted ring.

From these results it would appear that the restriction had the desired effect of increasing the magneto-transmission modulation depth, although the background transmission had been substantially reduced.

However, the processes occurring at the exit of the ring are more subtle than a simple blocking-out of the interference side-lobes in the interference pattern. The transmission of the output gap itself is reduced due to the increase in transverse energy the wavepacket is required to gain. For example, the original wavepacket had a forward energy component of 10meV and a transverse energy of 1.27meV giving a total of 11.27meV. The first sub-band energy in the gap region is 11.57meV, so that only the higher Fourier components of the wavepacket can find stable transverse states in the restricted region. If attention is focussed on just the forward component of the energy, then the effect of the increase in transverse energy of the wavepacket in travelling from the thicker wire to the thinner wire can be interpreted in terms of the wavepacket impinging on an electrostatic potential barrier. Indeed, if the conduction band for electrons is considered to be just the first transverse sub-band in each wire, then a thinning (or thickening) of the conductors causes a potential discontinuity in the conduction band similar to the energy discontinuity at a heterojunction.

In the simulation this effect manifests itself not only as a decrease in transmission, but the transmitted wavepacket in the constricted region is seen to move much more slowly because most of the energy has been converted into the transverse component.

Finally it is noted that this method of obtaining single-mode propagation at the exit of the ring may be preferable to thinning of the wire all round the ring as this would be likely to decrease the transmission through the input junction and thus increase the background resistance.

Figures 9.21 to 9.40: The effect of constricting the exit of the ring in the absence of an applied field.

Fig 9.21

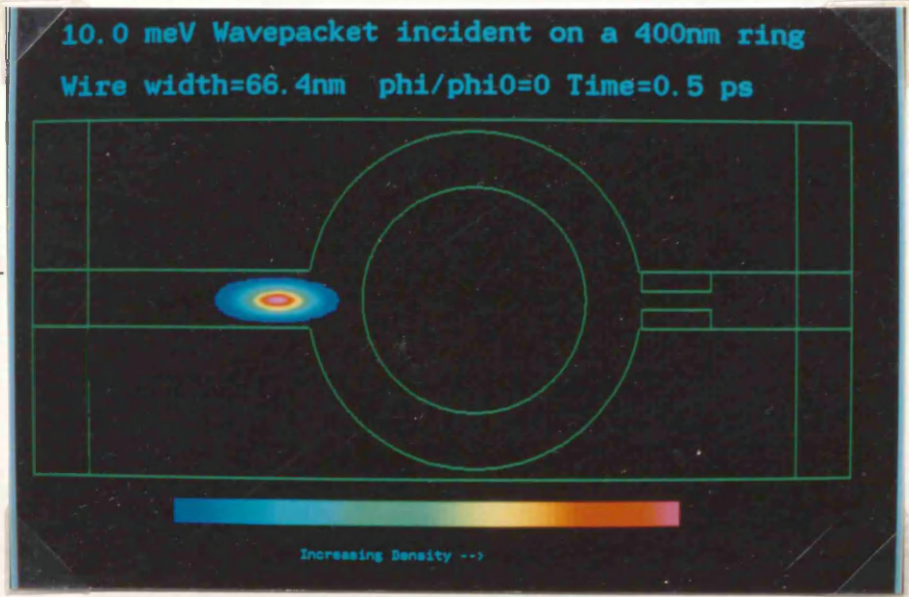


Fig 9.22

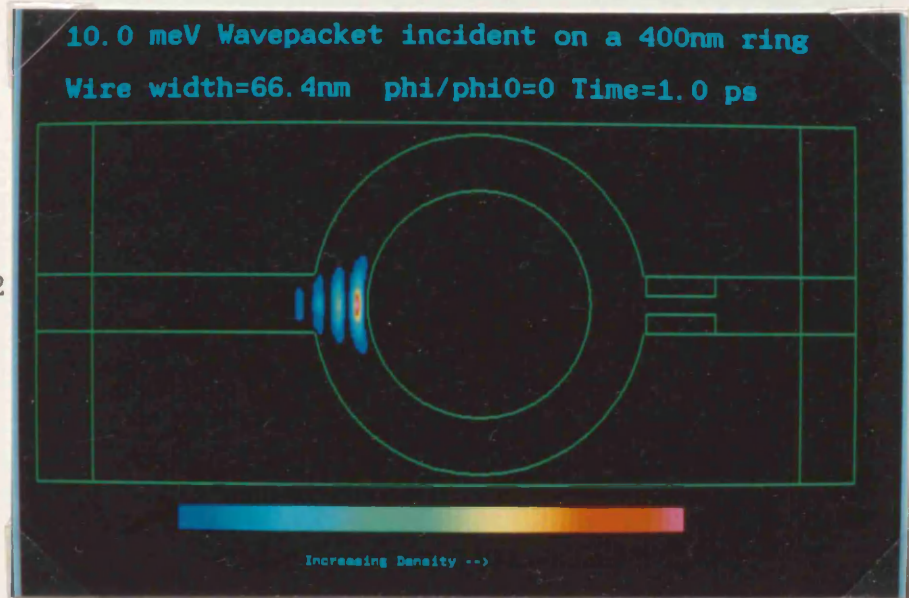


Fig 9.23

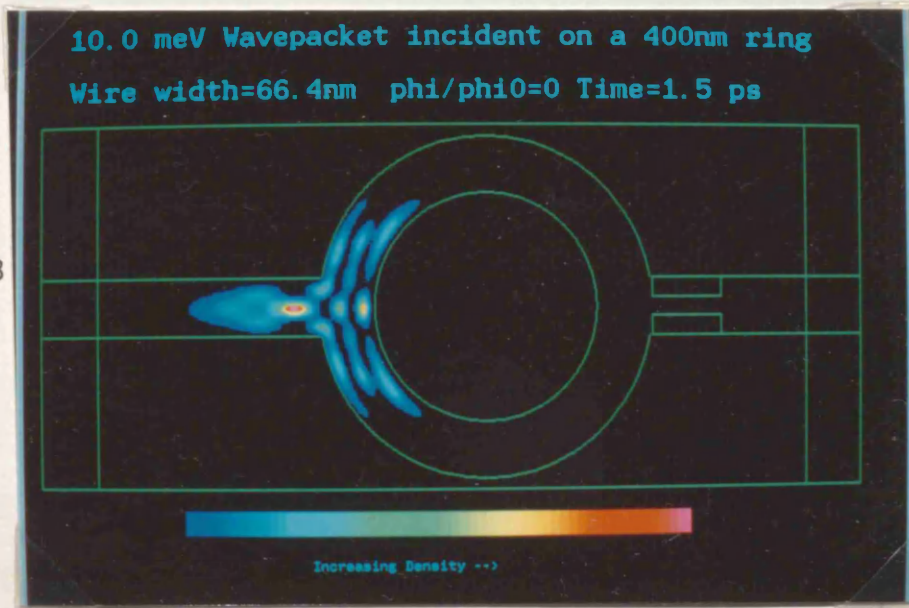


Fig 9.24

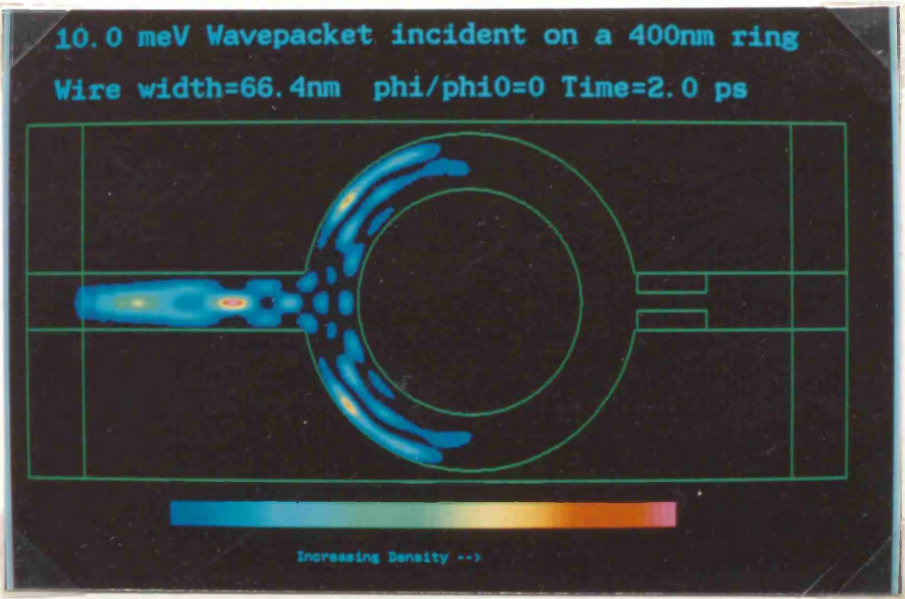


Fig 9.25

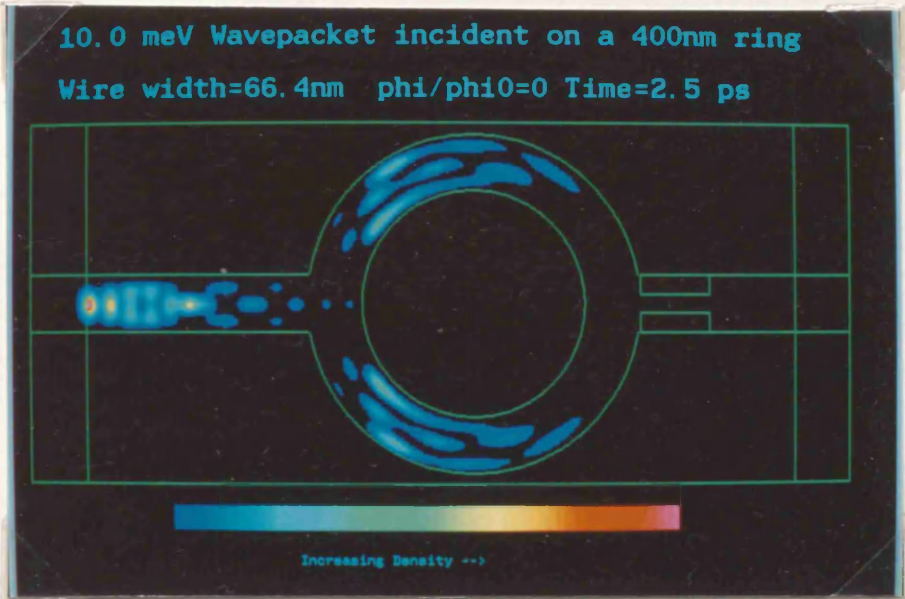


Fig 9.26

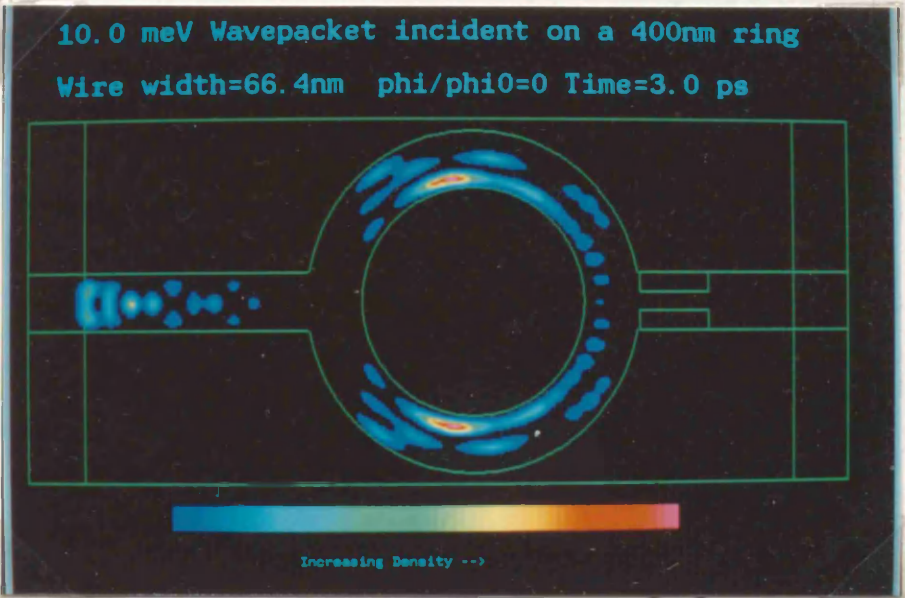


Fig 9.27

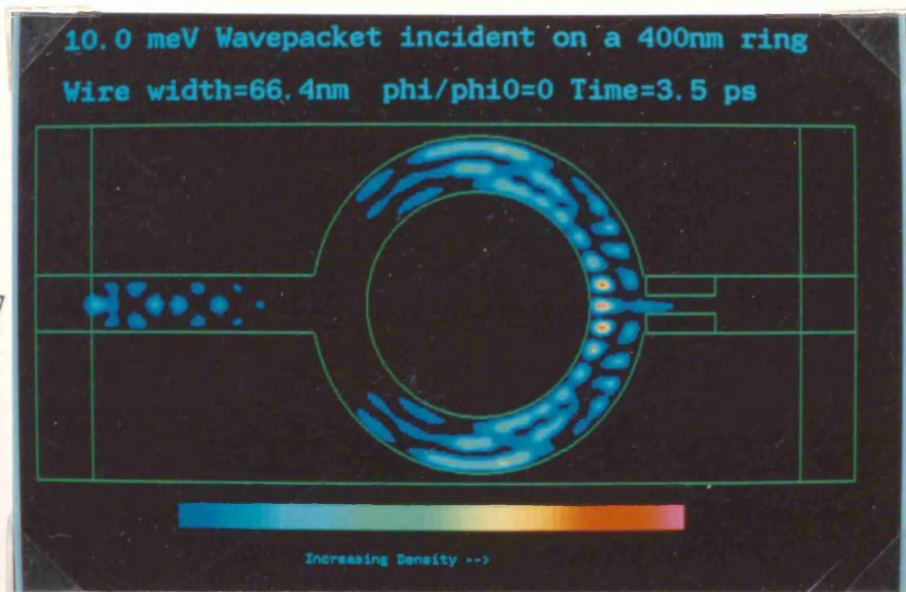


Fig 9.28

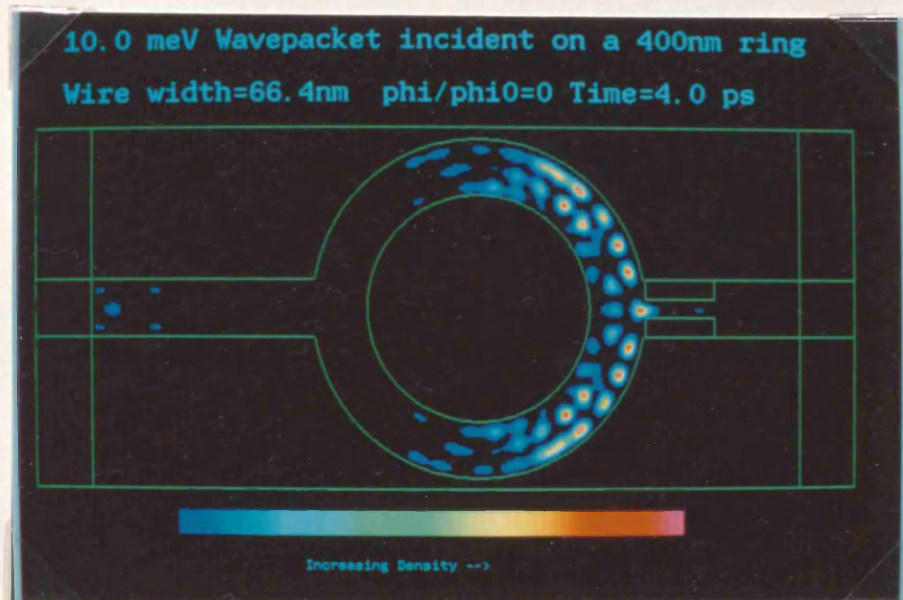


Fig 9.29

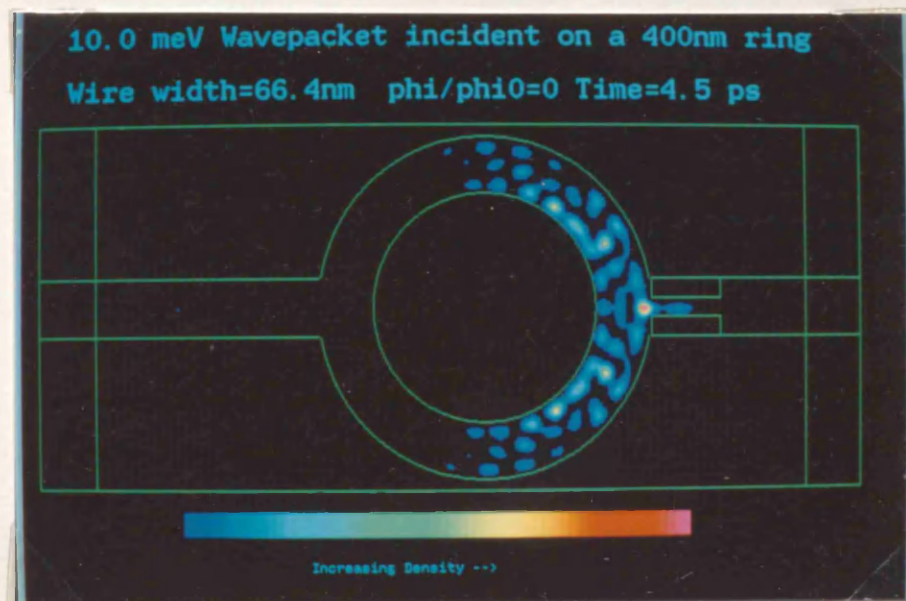


Fig 9.30

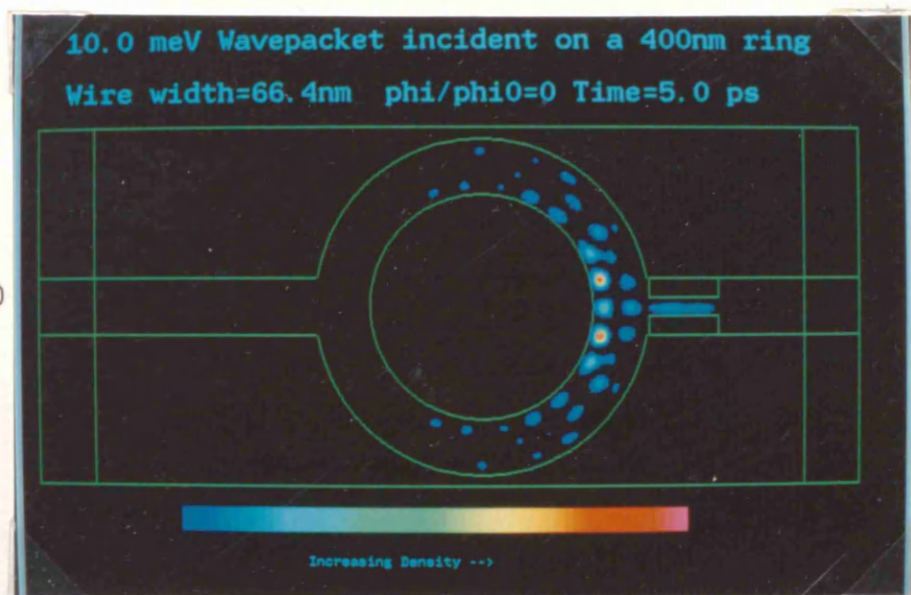


Fig 9.31

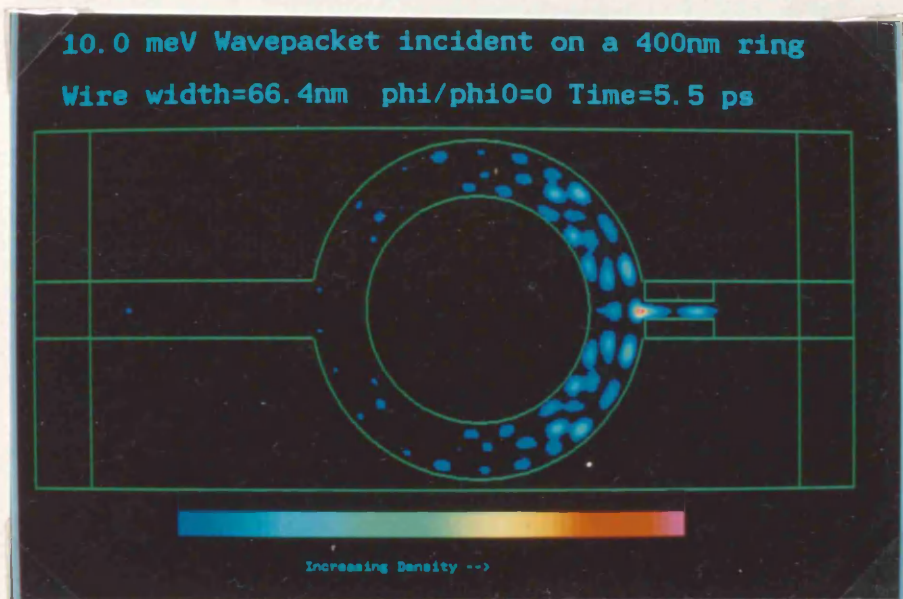


Fig 9.32

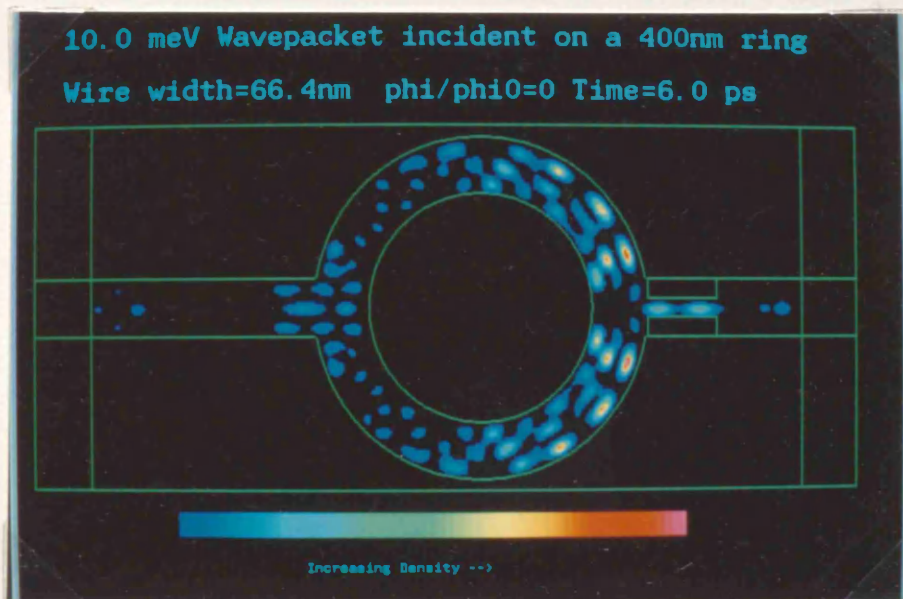


Fig 9.33

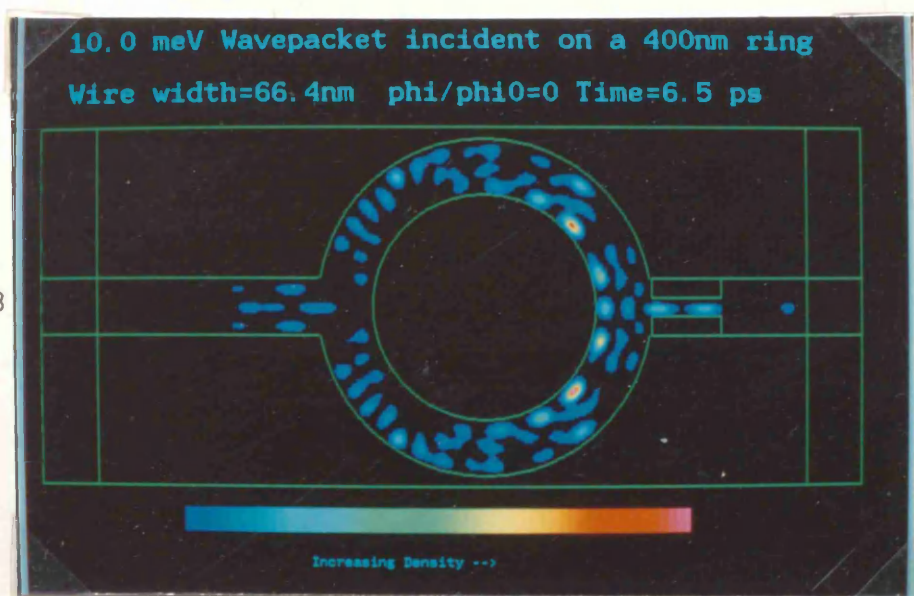


Fig 9.34

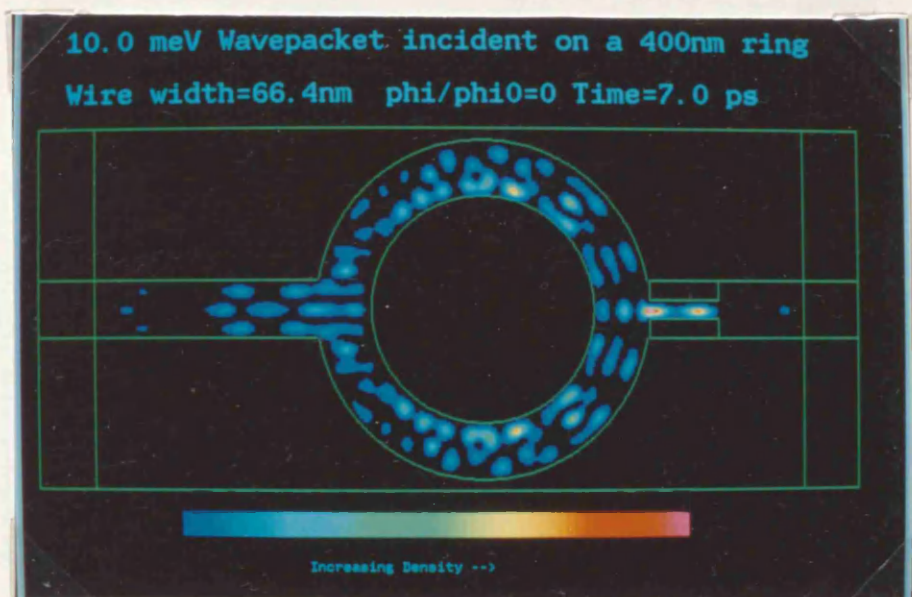


Fig 9.35

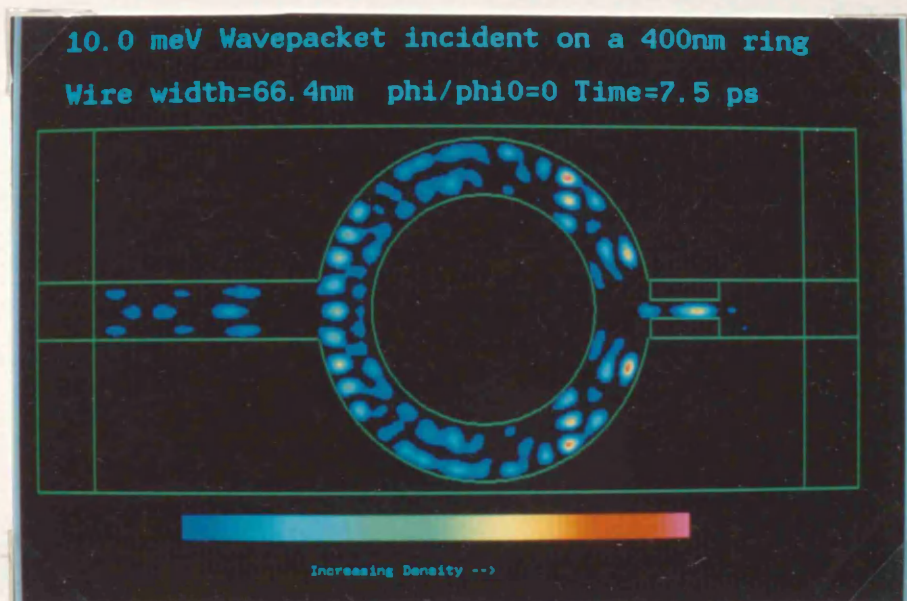


Fig 9.36

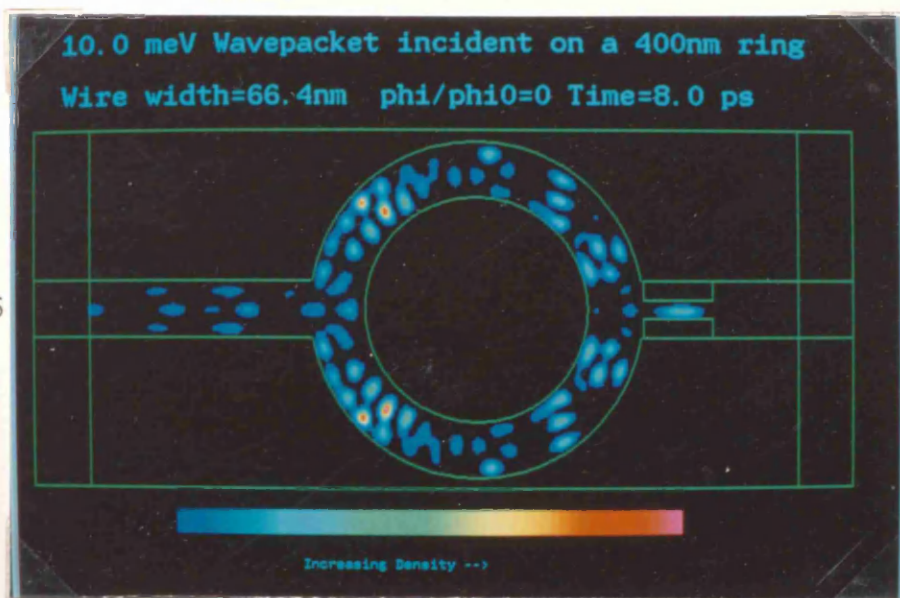


Fig 9.37

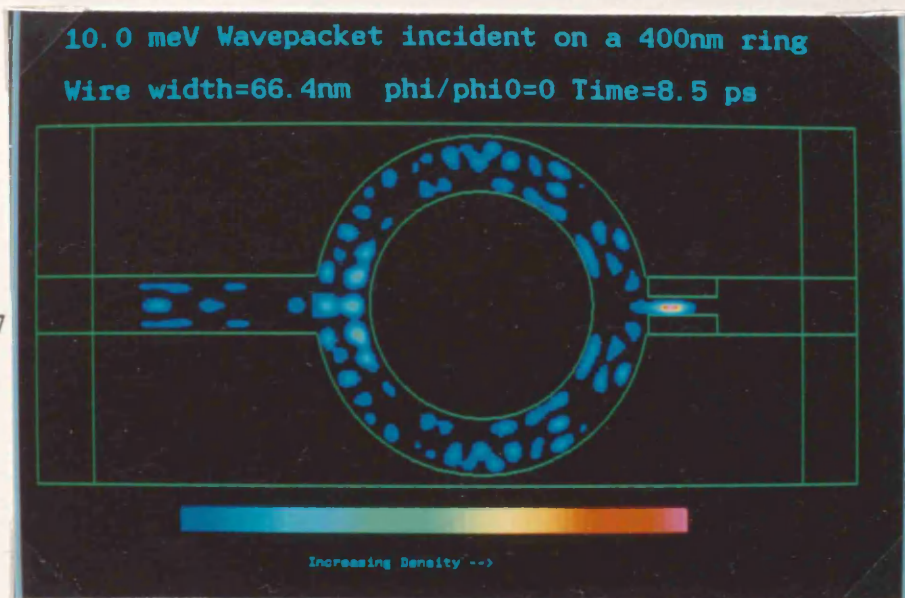


Fig 9.38

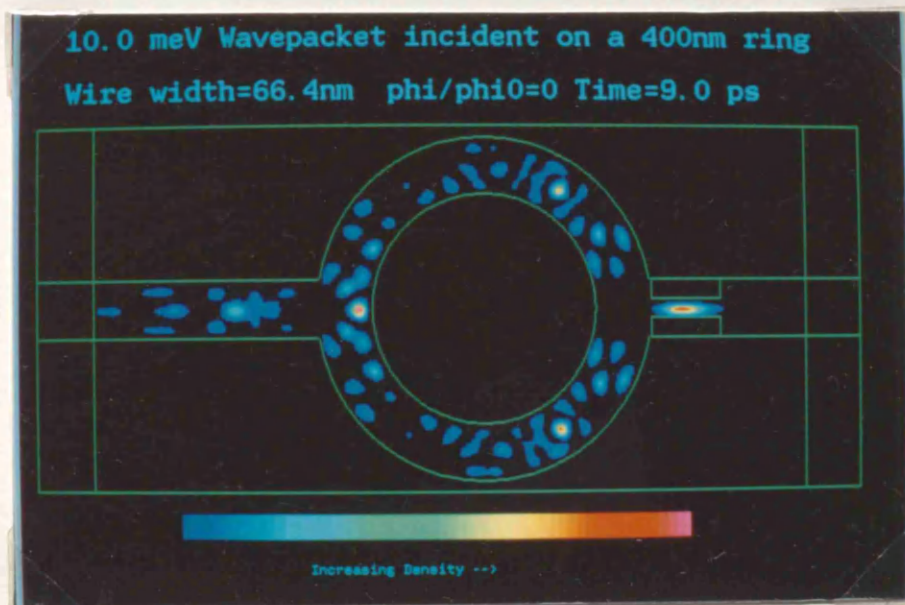


Fig 9.39

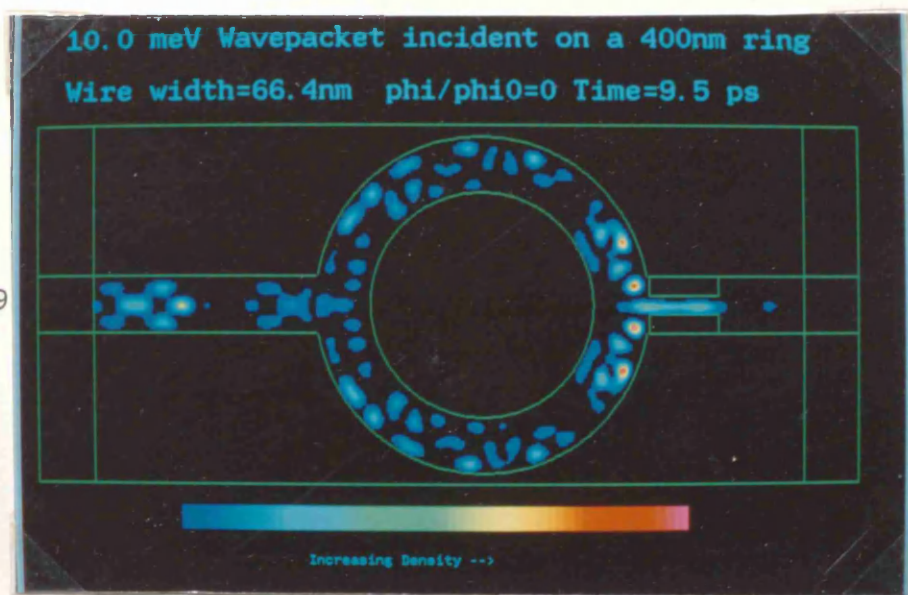
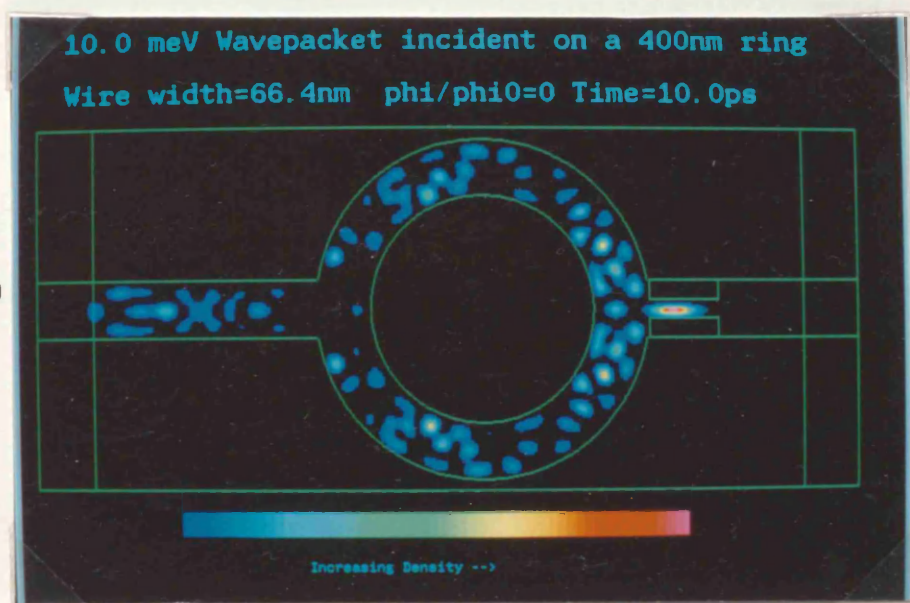


Fig 9.40



Figures 9.41 to 9.46: The effect of applying half a flux quantum to the constricted ring.

Fig 9.41

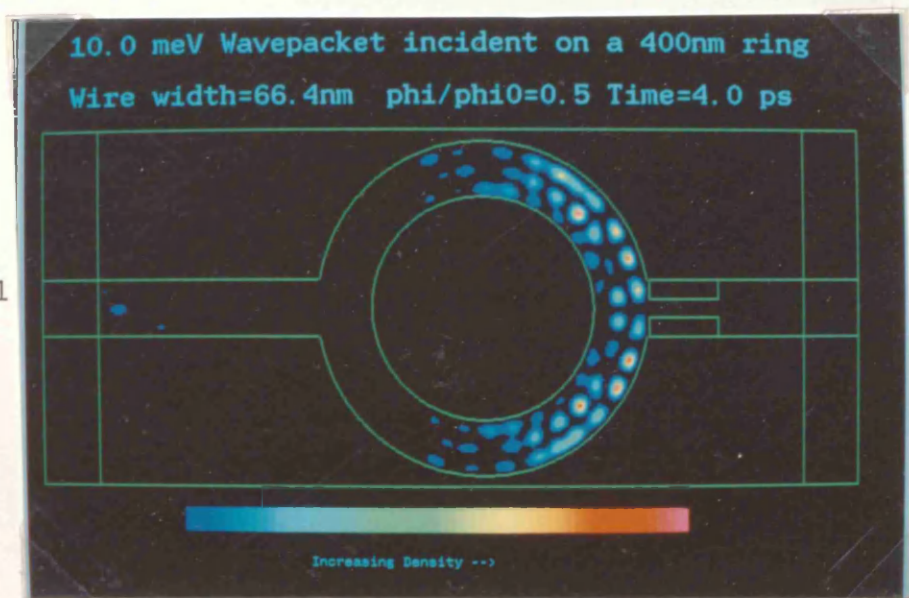


Fig 9.42

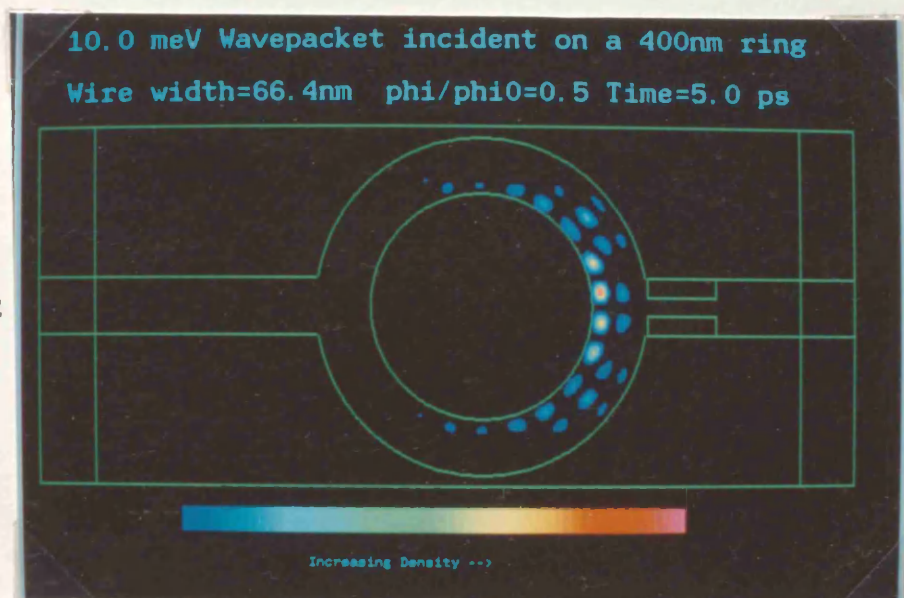


Fig 9.43

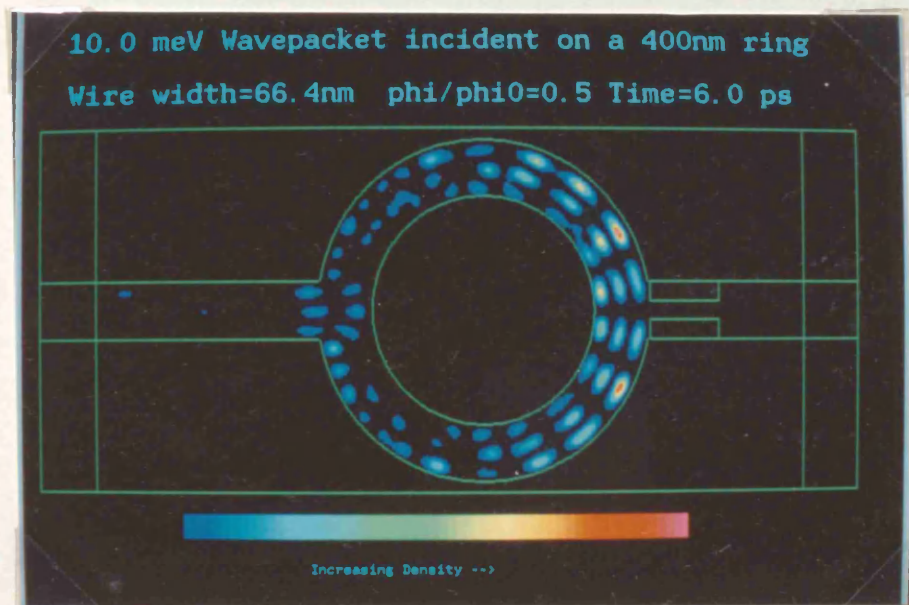


Fig 9.44

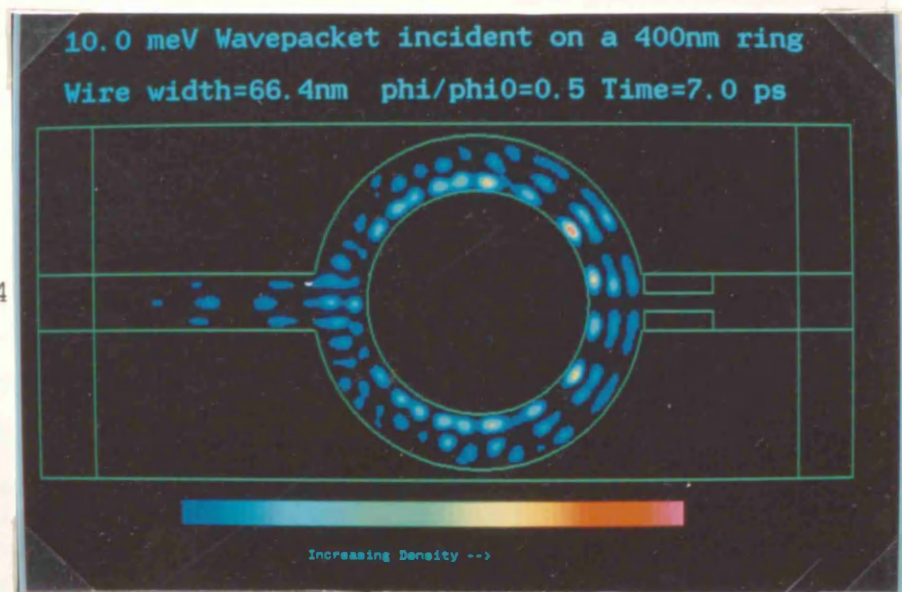


Fig 9.45

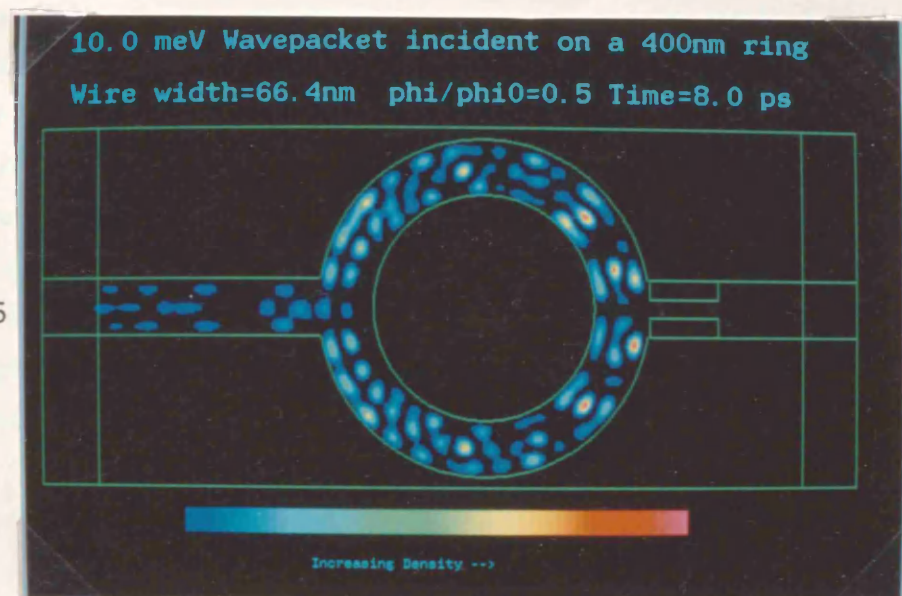
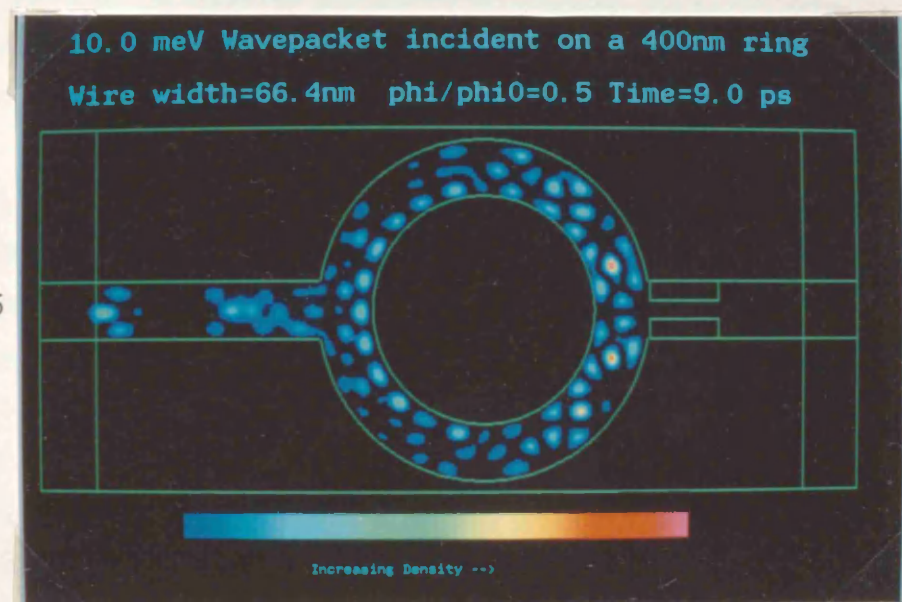


Fig 9.46



9.8 Other factors influencing the magneto-transmission characteristics of Aharonov-Bohm rings.

9.8.1 The effect of the fluctuation potential caused by remote donors.

Up to this point the computer model has dealt with an idealised structure with unperturbed confinement potentials. In reality, there will be fluctuations in the potential within the wires arising from the presence of the remote donors. This section seeks to explore qualitatively what effect this has on the magneto-resistance characteristics of the ring.

The number of ionised donors in the modelled device was estimated to be about 470 from the sheet electron concentration quoted experimentally. These donors were considered to be randomly distributed in three-dimensional space between 5 and 150nm above the plane of the 2DEG and the fluctuation potential was calculated at each mesh point by summing contributions from each donor assuming an unscreened Coulombic potential distribution. However, a straightforward calculation of this type taking only the 470 donors enclosed by the computational mesh, leads to an unrealistic accumulation of potential towards the centre of the grid. This is because a finite system is being modelled instead of a small section of an extended system for which the potential should be more evenly distributed.

To obtain a better approximation to the continuous system, extra neighbours outside the grid (but having the same random distribution as those inside) were considered. The potential was then calculated by considering the potential due to each charge inside the grid for each mesh point and also the potential due to the nearest (to the mesh point) neighbour replica outside the grid.

The potential arising from a number of random charge distributions was calculated. The results showed that the spatial scale of the fluctuation was relatively large, being about the same size as the ring itself. The fluctuation potential was only superimposed on the confinement potential in the region occupied by the ring only, leaving the input and output wires without any perturbation. This was done to avoid any problems in launching the initial wavepacket. To account for the fact that the wavepacket propagates into the perturbed region from a zero potential region, a constant potential was subtracted from the potential distribution so that the sum of the potentials over all points was zero ensuring that an evenly distributed wavefunction should not gain or

lose energy over a sufficiently large area. The interface between the perturbed and unperturbed regions was also smoothed by a Gaussian envelope to reduce unphysical reflections.

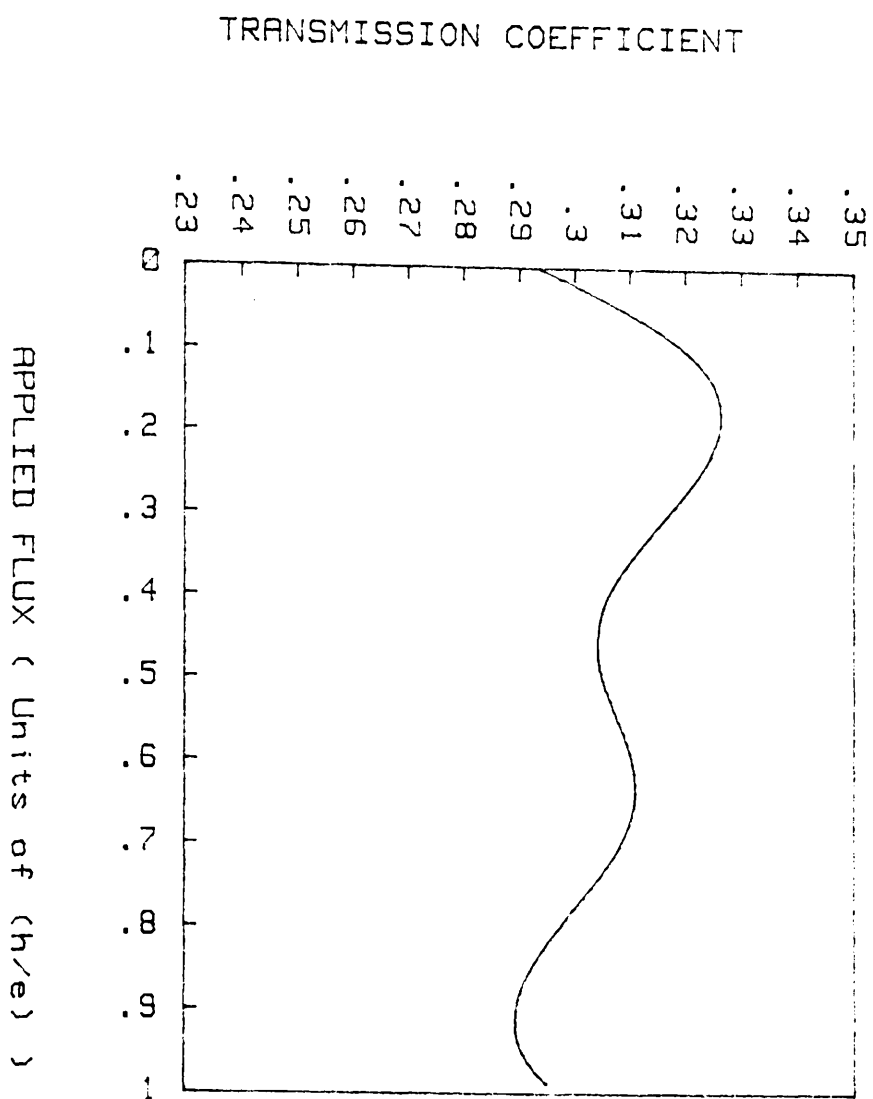
After subtracting the average potential, the r.m.s value of the fluctuation was about 30meV, which, being three times the average wavepacket energy, reduced the zero-field magneto-transmission of the ring to a very small value. Since it was known that such rings do conduct, it was concluded that the model was probably too crude in predicting the fluctuation amplitudes and/or that the procedure for launching the initial wavepacket was unrealistic. However, it was assumed that the general form of the calculated fluctuation potential was correct and it was only the amplitude of the fluctuation that was incorrect. The r.m.s amplitude of the fluctuation potential was therefore scaled so that the zero-field magneto-transmission was of the same order as the transmission without the fluctuation potential.

Figure 9.47 shows the magneto-transmission curve for a 10meV wavepacket propagating through a 400nm diameter ring when the r.m.s value of the fluctuation potential was 1.0meV. Although the average value of the transmission of about 0.3 is slightly lower than the same ring without the fluctuation potential present, the modulation depth of the oscillations has not been significantly changed. The positions of the main peaks and minima are also approximately in the same position as before. The main difference is that the small subsidiary peaks apparent in the magneto-transmission of the unperturbed ring are not seen in the perturbed case. It is thought that these smaller peaks are a result of multiple-transit trajectories and the smoothing of the magneto-transmission curve when the fluctuation potential is applied is probably a result of suppressing these trajectories. If the transmission coefficient of the arms of the ring are quite near unity, as the average value of the transmission suggests, the amplitude of multiple-transit partial waves could still be appreciably reduced by an additional factor (dependent on the arm transmission) on each pass around the ring.

To obtain more conclusive results, analysis of more magneto-transmission data for different impurity configurations would be necessary to show that any results were statistically significant and that the features of the curve obtained were not just specific to a singular case.

Figure 9.47

MAGNETO-TRANSMISSION OF A 400nm DIAMETER A-B RING



9.8.2 The effect of a longitudinal electric field.

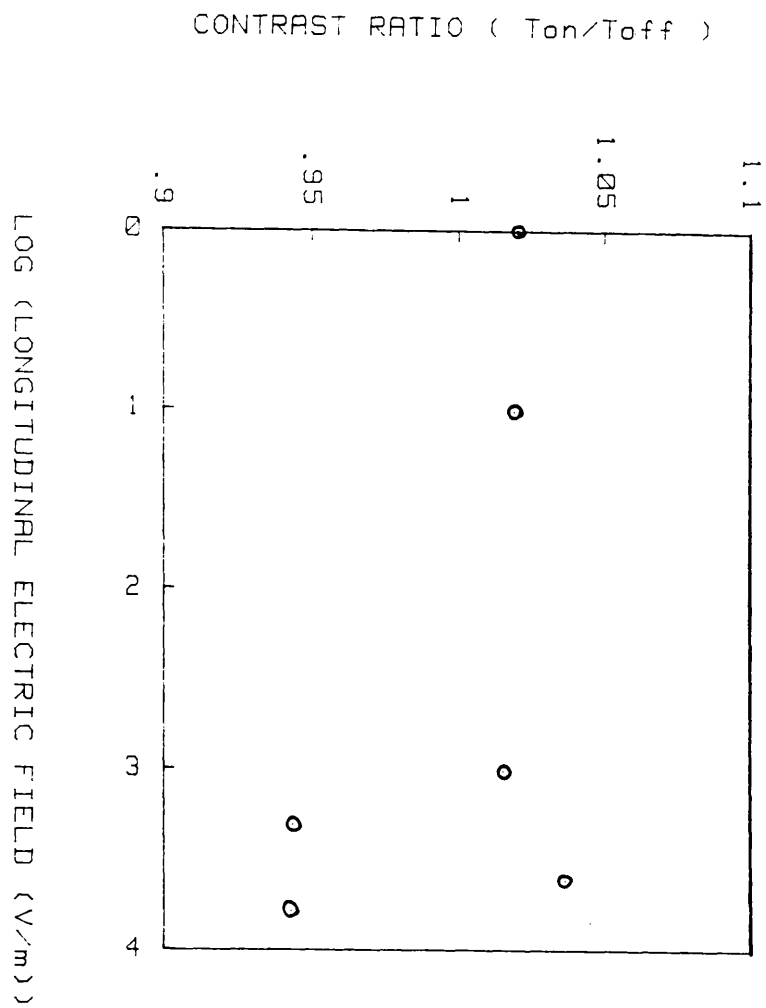
In all the calculations performed so far it has been assumed that there is no electric field across the device and that the wavepacket had acquired its energy from an electric field outside the region of interest. This was considered to be a good approximation to the experimental conditions as electric fields across the device are kept to a minimum to avoid carrier heating which would lead to an increase in scattering. Typical experimental fields are of the order of 1V/m [Timp *et al* 1987]. The computer model can however be used to test the effect of such an electric field explicitly. This was implemented by adding an extra potential energy term to the lumped potential $U(x,y)$ in the computer model (chapter 7).

It was not practical to undertake a calculation of the full magneto-transmission curves at any reasonable resolution for each electric field of interest in the time available. However, since the electric field is applied in the longitudinal direction, it was not expected to cause any asymmetries in the wavefunction and therefore the maximum and minimum transmission was expected to occur at approximately the same flux values as the zero-field curve. Therefore in order to assess the effect of the longitudinal electric field on the AB oscillation amplitude, the ratio between the transmissions at $\phi=0$ and $\phi=1/2(h/e)$ or "contrast ratio" was calculated. Figure 9.48 shows the results of these calculations for a 400nm ring (described in chapter 8) from which it can be seen that the contrast ratio remains close to unity under the influence of electric fields up to 7×10^3 V/m. The absolute values of the transmission and reflection coefficients did not differ greatly from the zero-field case, typically being in the range 0.33 to 0.39. It would therefore appear that typical experimental electric fields do not cause a significant change in the amplitude of the AB oscillations.

Considering the ballistic propagation of a wavepacket occupying the first transverse mode across the channel, one might expect sharp decreases in the transmission contrast ratio at certain threshold values of electric field. These would correspond to values at which the wavepacket would have acquired sufficient energy to transfer to the 2nd transverse sub-band or from the 2nd to the 3rd sub-band and so on. However, because the contrast ratio is so close to unity by the time 3 transverse sub-bands are occupied, the effect would probably only be significant for transfers up to the 3rd sub-band. This is clearly an area for further study.

Figure 9.48

VARIATION OF CONTRAST RATIO WITH ELECTRIC FIELD



9.9 A Fluid dynamical analogy.

9.9.1 General

This section outlines a fluid-dynamical analogy for the propagation of the wavefunction around an AB ring. It is discussed here because it offers an interesting and useful way of viewing some of the processes occurring and may be particularly useful in understanding behaviour in high magnetic fields.

Madelung [1926] has shown that the Schrodinger equation could be formulated in terms of two fluid-dynamical equations. These are obtained by writing the wavefunction in the following form:-

$$\psi = \rho \exp(i\phi) \quad (\text{Eqn 9.11})$$

where ρ and ϕ are real variables.

Substituting this form into the time-dependent Schrodinger equation and equating the imaginary parts results in the flux continuity equation [Olariu and Popescu 1985]:-

$$d(\rho^2)/dt + \text{Div} ((\hbar \nabla \phi - e\mathbf{A}).\rho^2/m) = 0 \quad (\text{Eqn 9.12})$$

The variable ρ^2 thus appears as a density of the "fluid", as expected from the fact that the probability density is:-

$$\psi\psi^* = \rho^2 \quad (\text{Eqn 9.13})$$

Equating the real components results in an equation expressing energy conservation within the fluid:-

$$\hbar \frac{d\phi}{dt} + (1/2m) (\hbar \nabla \phi - e\mathbf{A})^2 + eV - (\hbar^2/2m) \nabla^2 \rho / \rho = 0 \quad (\text{Eqn 9.14})$$

The classical limit is obtained by letting $\hbar \rightarrow 0$ so that the quantity $(\hbar^2/2m) \nabla^2 \rho / \rho$ becomes negligible and the above equation is then seen to reduce to the classical Hamilton-Jacobi equation for a charged particle in an electromagnetic field [Goldstein 1959].

By comparison with the Hamilton-Jacobi equation a velocity field can be defined as:-

$$\mathbf{v} = (\hbar \nabla \phi - e\mathbf{A})/m \quad (\text{Eqn 9.15})$$

Since this thesis is concerned with the effects of magnetic fluxes on the wavefunction, it is of particular interest to find how magnetic fields are represented in the fluid model. This can be done by considering the vorticity of the velocity field, Γ .

$$\Gamma = \oint \mathbf{v} \cdot d\mathbf{r} \quad (\text{Eqn 9.16})$$

Which can be converted to a surface integral over the ring.

$$\Gamma = \oint \nabla \times \mathbf{v} \cdot d\mathbf{s} = (-e/m^*) \Phi \quad (\text{Eqn 9.17})$$

Thus a magnetic flux causes a vortex in the fluid and can be expected to cause turbulence in the quantum fluid flow at high fields. This model, then, could offer a useful interpretation of the behaviour of the wavefunction at high magnetic fields.

9.9.2 The role of an imaginary potential in the fluid model.

The effect of the imaginary potential introduced in section 7.8 can also be assessed using the fluid model.

Starting from the time-dependent Schrodinger equation and assuming that the magnetic field applied to the system is zero one can use the model Hamiltonian of section 7.8 to obtain:-

$$\begin{aligned} i\hbar (\psi^* d\psi/dt - \psi d\psi^*/dt) &= (-\hbar^2/2m^*) (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*) \\ &\quad + \psi \psi^* (V_r + iV_i - (V_r - iV_i)) \end{aligned} \quad (\text{Eqn 9.18})$$

which can be written as:-

$$i\hbar d(\psi \psi^*)/dt = (-\hbar^2/2m^*) \nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*) + 2iV_i \psi \psi^* \quad (\text{Eqn 9.19})$$

Identifying the probability current as:-

$$J = (\hbar/2mi) (\psi^* \nabla \psi - \psi \nabla \psi^*) \quad (\text{Eqn 9.20})$$

and writing $\psi = \rho \cdot \exp(i\phi)$ as before then yields the modified fluid continuity equation:-

$$d(\rho^2)/dt + \nabla J - 2V_i \rho^2/\hbar = 0 \quad (\text{Eqn 9.21})$$

This is seen to be of the same form as the normal continuity equation apart from the appearance of the extra term involving the imaginary potential. The minus sign indicates that this term can be thought of as a sink for particles, which is in agreement with the analysis of section 7.8 and the results of chapter 8.

9.9.3 Consequences of the quantum potential in narrow conductors.

Another interesting situation arises when the fluid equations for a restricted channel are considered. In fluid dynamics, the velocity of the fluid increases through the restriction in order to keep the volume of flow at each point continuous. Bernoulli's formula then shows that the increase in velocity is accompanied by a decrease in the pressure of the fluid, or in other words that energy is transformed from a pressure to kinetic energy. In the quantum-hydrodynamical formulation, the velocity is seen to *decrease* through the restriction. The reason for this may be seen by rewriting the energy conservation equation as:-

$$\hbar d\Phi/dt + (\hbar \nabla \Phi - eA)^2 + e(V + V_q) = 0 \quad (\text{Eqn 9.22})$$

where V_q is a "quantum potential" defined by:-

$$eV_q = -\hbar^2/2m^* (\nabla^2 \rho/\rho) \quad (\text{Eqn 9.23})$$

The quantum potential term can thus be interpreted as an additional potential arising from the Pauli exclusion principle. This can be seen from the functional form $-\nabla^2 \rho$ appearing, which is representative of the degree of confinement of the wavefunction.

By normalising this form to the density at each point, the quantity $-\nabla^2\rho/\rho$ can be considered to be the energy density due to the confinement of the wavefunction.

If a wavefunction in a thin wire is represented as a plane-wave in the x-direction modulated by the ground-state eigenfunction across the wire, the quantum potential for electrons moving in the x-direction can be found by assuming a separable wavefunction of the form:-

$$\psi(x,y) = \psi(x)\psi(y) = A\exp(ik_x x) \sin(k_y y) \quad (\text{Eqn 9.24})$$

where A is a normalisation constant.

By also separating the density variable $\rho(x,y)$, and equating the components to the corresponding wavefunctions, one can write:-

$$\rho(y) = A \sin(k_y y) \quad (\text{Eqn 9.25})$$

$$\rho(x) = 1 \quad (\text{Eqn 9.26})$$

The quantum potential for this system is then given by:-

$$eV_q = +\hbar^2 K_y^2 / (2m^*) \quad (\text{Eqn 9.27})$$

which is equivalent to the energy due to the transverse confinement of the wavefunction across the wire. This means that when the "quantum fluid" passes through a restriction the quantum potential arising from the spatial confinement of the wavefunction increases and leads to a *decrease* in the velocity as consequence of energy conservation. This can be seen in the simulations of wavepacket propagation through a ring with a constricted exit.

Note that this effect is purely quantum mechanical in nature and is not observed in conventional fluids as V_q is proportional to \hbar^2 and is therefore negligible in the classical limit.

9.10 Summary.

This chapter has presented the results of magneto-transmission calculations of wavepackets incident on AB rings.

The magneto-transmission characteristic a 400nm diameter ring was calculated using the two-dimensional time-dependent model developed in earlier chapters. The transmission coefficient of the ring was found by time integration of the currents passing into the left and right hand absorbing contacts.

The relationship between the transmission coefficient of the ring was discussed and the single-channel Landauer formula was used to obtain the resistance of the ring as a function of magnetic field.

The magneto-transmission characteristics of a ring fabricated by Ford *et al* [1987] were calculated for two different wavepacket energies. The results showed that the magneto-transmission characteristic is sensitive to wavepacket energy and the curve for the 9meV wavepacket showed a dip in the transmission at a flux value of $1/2(h/e)$ which was sharper than expected.

It was proposed that the structure apparent in the magneto-transmission characteristics were a result of resonant effects similar to those seen in the one-dimensional analysis of chapter 5. It was found that the shape of the magneto-transmission curves could be well approximated by assuming that the simulated transmission could be represented as a hybrid of the 1D model of chapter 5 and the unconstrained (but still one-dimensional) simple two-slit model of chapter 3.

The conductor width of the first ring manufactured by Ford *et al* was initially taken to be 80nm and a simulation of this ring gave a predicted magneto-resistance oscillation amplitude of between 20 and 30%. This was much higher than the experimental value of about 5%. However, there was later a good body of evidence to suggest that the value of 80nm for the conductor width was an under-estimate and that the true value was probably around 150-200nm.

Simulations using this revised width of 200nm yielded a magneto-resistance oscillation amplitude of about 12% which was in better agreement with experiment especially considering the remaining uncertainty in the precise conductor width.

A ring of different geometry was fabricated by Ford *et al* [1989] in which the width, estimated from magnetic depopulation studies, was thought to be about 80-100nm at the highest gate bias. The measured oscillation amplitude for this ring reached a maximum value of about 20% of the background resistance which was in better agreement with the earlier simulation of a ring with 80nm wires. These results support the conclusion that the number of transverse modes across the exit of the ring is a major factor controlling the amplitude of the AB oscillations.

Based on this hypothesis, a modified ring structure was simulated with the objective of increasing the AB oscillation amplitude. This was done by constricting the exit of the ring so that only the central mode could propagate along the output wire. The results showed that the ratio of the on-to-off transmission could be improved from about 1.04 to about 8.8 .

Other factors that might influence the magneto-transmission characteristics were then investigated. The effects of random potentials arising from the remote ionised donors were assessed and it was concluded that if the ring was capable of conduction at all, the amplitude of the magneto-resistance oscillations would not be not be significantly changed. What was significantly changed was the general form of the oscillations, which indicated that the high-frequency components of the fluctuation were appreciably suppressed. The effect of a longitudinal electric field was also investigated and it was found that typical electric fields present in experiment would have very little effect on the magneto-resistance oscillation amplitude (provided that transport remained ballistic).

Finally, a fluid dynamical model of wavepacket propagation in these structures was described. This model offers an alternative view of the role of magnetic fluxes in AB devices and of ballistic transport through constricted channels.

CONCLUSION OF THESIS.

This thesis is the first study of quantum ballistic transport in a two-dimensional quantum waveguide network. The theory developed was tested by comparing calculated against measured magneto-resistance oscillation amplitudes for an AB ring formed from the 2DEG at a heterojunction interface. Good agreement with experiment was obtained and the method can thus be considered a major improvement on earlier one-dimensional models.

The study was made possible by the development of a two-dimensional time-dependent Schrodinger equation solver, based on the ADI method, which incorporated two types of magnetic field distribution and which was also able to successfully model the contact regions at each end of the computational mesh.

A large part of the thesis was therefore devoted to the numerical techniques necessary for the efficient solution of the two-dimensional problem and the method chosen offered considerable savings in computation time over eigenfunction expansion techniques. The essential step in this method was to resolve the Hamiltonian operator into two components on which one-dimensional solution strategies could be used. The consequences of decomposing the Hamiltonian were investigated and it was found that the error generated in this procedure was proportional to the magnetic field and this was thought to be the reason behind the slight numerical instability present at high magnetic fields. This was a disappointment since a study of wavepacket propagation in ring structures at high magnetic fields would have been of interest in assessing a possible damping of the resistance oscillations. An estimate of the behaviour of the transmission coefficient at high fields based on an extrapolation from the low field results was given instead. There is, however, scope for further work to improve the stability of the algorithm which could be approached by considering the commutators of the two parts of the Hamiltonian.

The problem was ultimately reduced to the solution of several sets of tri-diagonal matrix equations for which standard linear algebra techniques could be used. However, in order to obtain the best performance from the vector-processing facility on the IBM3090, a non-standard re-organisation of the algorithm was necessary to prevent recursions inhibiting vectorisation of the code. This was of major importance because of the very numerically intensive nature of the calculation.

The full two-dimensional study was preceded by a time-independent study of a ring with one-dimensional conductors. These studies formed the background work for the project and were later used for comparison with the two-dimensional results. The effect of confining the electrons in rings composed of one-dimensional conductors was to introduce an energy dependence into the expression for the transmission coefficient of the ring, making the behaviour considerably more complex than a simple periodic function of the enclosed flux predicted by an elementary model. Two basic types of behaviour were found; that when the ring was poorly coupled to its external leads and that when it was well coupled.

In the poorly-coupled regime, the peaks in the transmission characteristics could be understood by considering the limiting case to be modelled by a freely orbiting electron for which the energy eigenvalues are well known.

The well-coupled regime with no scattering included in the arms of the ring was representative of the idealised structures which were being considered, and although the peaks and flux-dependence of the transmission could be understood from the work of Buttiker *et al* [1985], the origin of the sharp drops in transmission was uncertain. To explain this behaviour, a phasor diagram approach was taken in which it was shown that even if the magnitude of the wavefunction at the exit to the ring was not zero the transmission could still be zero if the resultant wavefunction at the was entirely real so that no current could flow. These regions can therefore be thought of as defining the Brillouin zones for the allowed energy bands of the ring.

Two-dimensional transport in AB rings was initially studied by using the Schrodinger equation solver to simulate the time evolution of a Gaussian wavepacket through a ring structure with conductors of a definite width. The results from these time-dependent studies can be summarised as follows.

- 1 The AB effect occurs in a fully two-dimensional ring and the transmission can be reduced to virtually zero by the application of half a flux quantum through the hole in the ring provided that the conducting channels are narrow enough to allow only single-mode propagation.

- 2 Increasing the size of the system (so that there is more than one transverse mode) resulted in a great increase in the complexity of the probability density distribution.

3 Increasing the width of the conducting channels resulted in the formation of a multi-mode structure in the probability density (in this case three modes were observed). The effect of having more than one mode present is to decrease the magneto-transmission oscillation amplitude.

4 The direct action of the magnetic field on the wavefunction in the conducting channel was to displace the whole of the interference pattern. This mechanism could lead to the damping of the magneto-transmission oscillations at sufficiently high fields (estimated to be about 0.5T for this ring).

Although the time-dependent simulations provided a valuable qualitative picture of the nature of wavepacket propagation through these structures, it was important to test the theory by quantifying results and comparing with experiment. To do this it was necessary to find the transmission coefficient of the ring as a function of magnetic field from which the magneto-resistance characteristics could be determined. The time-dependent model was again used to propagate the wavepacket until interaction with the ring was judged complete. The resistance of the ring was then calculated from the asymptotic transmission coefficient using the single-channel low-temperature Landauer formula which was shown to be valid under the conditions considered.

Initial magneto-resistance calculations on a ring considered earlier for the time-dependent studies showed that the magneto-transmission modulation depth was much less than 100% predicted by one-dimensional models.

The key concept to emerge was that the modelled ring can be thought of as being intermediate between two extreme cases, that in which the conducting channels are strictly one-dimensional and that in which the conducting channels are infinitely wide. In the one-dimensional case, complete modulation of the current as a function of the enclosed flux can be obtained but only minimal information concerning the interference pattern at the ring exit can be gathered by the external leads. At the other extreme, a complete knowledge of the interference pattern can be obtained but without any modulation in current.

The magneto-transmission data were therefore analysed in terms of a hybrid 1D and open-geometry model. This gave a good approximation to the shape of the simulated magneto-transmission curves and also provided a method for obtaining an estimate of the parameters needed for the one-dimensional model.

The theory was then compared with experiment by simulation of the magneto-resistance characteristic for a ring fabricated by Ford *et al* [1987]. The first results produced a magneto-resistance oscillation amplitude of about 20-30% of the background value which was rather higher than the experimentally obtained value of about 5%. However, it was later shown that a major contributor to this discrepancy was the uncertainty in the conductor widths present in real devices. This remained a serious problem in preventing more accurate comparisons between theory and experiment.

The conclusion drawn from these results was that the transmission of the side-lobes of the interference pattern at the exit of the ring was responsible for the loss in the magneto-resistance modulation depth and that a narrow channel width is essential if large AB oscillations are to be seen.

Another problem with the simulation of the magneto-transmission curves was the amount of computer time required. This was because the wavepacket had to be propagated for a relatively long time to obtain the asymptotic transmission coefficient. This had to be repeated for each point on the magneto-transmission curve for each magnetic field value and although it is the magneto-resistance curves which are most useful for comparison with experiment, the number of magneto-resistance calculations that could be performed were necessarily limited.

The root of the problem is that a time-dependent method had been used to find a solution to a problem which would be better solved using time-independent methods. One method which is becoming increasingly attractive in this respect is a two-dimensional network solution method which is currently being developed by Pepin [1989].

Recent theoretical work shows that the two-dimensional wavefunction can be represented in a one-dimensional form on a mesh that covers the whole region of interest. Providing one knows the appropriate S-matrix for each of the 4-port mesh points, the outgoing plane-waves from the system can, in principle, be found from a knowledge of the incoming waves. This is done by solving a system of equations involving the S-matrix elements for each of the nodes on the mesh and the appropriate T-matrix elements describing the propagation between the nodes. To make this method viable for large systems, the construction of the network equations must be automated, and a computer program which does this is currently under development.

This time-independent method can be expected to be quicker than the present method and would enable a greater number of h/e cycles in the magneto-resistance characteristic to be calculated and make the Fourier transforms more meaningful. Provided the direct action of the fields could be incorporated into this model, the high-field magneto-resistance behaviour could also be investigated. The advantage of calculating the transmission versus energy curve for the ring would be that it would only have to be calculated once. Transmissions of different shaped wavepackets could then simply be investigated by Fourier analysis of the wavepacket to find the appropriate plane-wave components.

If further progress in time-independent methods is not forthcoming, progress can still be made with the time-dependent model. This would be desirable in any case as it would enable more accurate calculations to be performed in the time domain. Although the numerical algorithms have been optimised as far as can reasonably be expected on present-day serial computers, there exists considerable opportunities for the implementation of the algorithms on parallel machines. Since each row (or column) in the two-dimensional mesh is solved independently for each half-sweep, the algorithm would be an ideal candidate for implementation on a transputer system.

Although in this thesis no attempt has been made at self-consistently solving Poisson's and Schrodinger's equation, this could be important future work in view of the fact that the magnetic fields cause asymmetries in the charge distribution. This in turn may then cause significant changes in the phase of the wavefunction through the electric-potential part of the four-dimensional electromagnetic flux.

A major disappointment with these AB devices is that they only produce oscillations of observable amplitude at very low temperatures, typically around 100mK. The currently favoured theory is that there is some energy bandwidth or "correlation range" E_C in which separate electron waves generate spatially correlated interference patterns over the region of interest. For low temperatures this bandwidth is approximately $K_B T$ and can be related to the length of sample L , over which interference patterns are correlated via [Stone and Imry 1986]:-

$$E_C = \hbar V_f / L$$

where V_f is the Fermi velocity.

In order to obtain a correlated interference pattern for the patterned-gate HEMT device considered in this work a temperature of the order of 10K would be required. This figure however is still about two orders of magnitude higher than those required experimentally to observe oscillations in micron-sized rings. This contrasts sharply with the negative differential resistance behaviour found in resonant tunnelling experiments which has now been observed at room temperature [Razeghi *et al* 1987]. It would be tempting to ascribe this to the smaller dimensions of resonant tunnelling devices, but this is probably not a complete explanation. The AB effect is critically dependent on the spatial interference pattern produced by electron waves which may have a certain energy spread whereas tunnel devices do not depend on any specific spatial correlation between electron waves of different energies. Furthermore the resonant peak in the transmission coefficient of a resonant double-barrier potential serves as an energy filter for electrons. Another aspect of conduction in the 2DEG is that electron concentrations are generally very high and little is known at present about what effect many-body interactions might have on the ballistic transport of electrons through these types of structures.

In overall conclusion, this thesis has made the first steps towards the modelling quantum waveguide networks. Considering the experimental uncertainty in determining conductor widths and the apparent sensitivity of the AB oscillation amplitude to the width, good agreement with experiment has been achieved.

Although much is being learnt about quantum conduction processes in these structures, prospects for the engineering applications of these devices do not look promising at present due to the very low temperatures required. The processes controlling the length and temperature scales over which coherent interference phenomena are observed would therefore be an important objective for future studies. Although theories do exist [Stone and Imry 1986] [Washburn and Webb 1986] concerning this subject, there is still a large amount of uncertainty particularly in two-dimensional systems.

From an engineering point of view, progress towards higher temperature operation may be achieved by considering these AB rings to be components in some kind of hybrid device. For instance, if the energy bandwidth of incident electrons is responsible for the very low temperatures required, combination with resonant tunnelling devices to act as energy filters might be considered. Possibilities of this kind could be initially assessed theoretically using the techniques developed in this thesis.

Finally, the techniques developed in this thesis could also be used to model quantum ballistic transport in other types of structures. In particular, the study of modulated-width quasi-1D wires promises to be an interesting area for future research. In these structures effective potentials can be achieved in the 1D conduction sub-bands by a narrowing of the conducting channel and the same type of behaviour that was present at the exit of the constricted ring in section 9.7 can be expected. Theoretical progress being made in this field using a one-dimensional coupled-mode theory by Barker and Laughton [1987].

Appendix A: The transmission matrix.

This appendix describes the details of the transmission matrix technique for calculating the transmission coefficient of a potential barrier. The transmission matrix is defined, and its properties (arising from current conservation) are derived. The details of the calculation for a model potential barrier are also given.

Definition.

Consider an arbitrary one-dimensional potential $V(x)$ depicted in figure A1, which is constant outside the domain $-l < x < l$. It is assumed that the barrier is centred on $x=0$ as this simplifies calculations, but this requirement is not essential.

One can then write down the wavefunction for a free particle either side of the interval in terms of its momentum eigenfunctions. It is assumed initially that the forms of the wavefunctions are valid for very large distances away from the barrier, so that the particles are truly free. Thus it is implied that in calculations of this type the source of particles giving rise to the incoming plane wave is effectively located at $x \rightarrow -\infty$. The two independent solutions to the free-particle Schrodinger equation to the left of the domain are:-

$$\psi_{1k}(x) = \exp(ikx) \quad (\text{Eqn A1})$$

$$\psi_{2k}(x) = \exp(-ikx) \quad (\text{Eqn A2})$$

the general form of the wavefunction in this region being some linear combination of the two independent solutions. Consider first the ψ_1 component. In order that the wavefunction to the right of the region be different (in general) from the solution on the left, the solution on the right must also be some linear combination of the two solutions $\exp(ikx)$ and $\exp(-ikx)$. Using a notation for the linear coefficients which anticipates reduction to matrix form, the ψ_1 component behaviour either side of the barrier can be written as:-

$$\psi_{1k}(x) = \exp(ikx), \quad x < -l \quad (\text{Eqn A3})$$

$$\psi_{1k}(x) = T_{11} \exp(ikx) + T_{21} \exp(-ikx), \quad x > +l \quad (\text{Eqn A4})$$

Similarly for the ψ_2 component:-

$$\psi_2(x) = \exp(-ikx) \quad , \quad x < -l \quad (\text{Eqn A5})$$

$$\psi_2(x) = T_{12}\exp(ikx) + T_{22}\exp(-ikx) \quad , \quad x > +l \quad (\text{Eqn A6})$$

The most general result for the wavefunction to the left and right of the region is:-

$$\psi_k(x) = A.\exp(ikx) + A'.\exp(-ikx) \quad , \quad x < -l \quad (\text{Eqn A7})$$

$$\psi_k(x) = C.\exp(ikx) + C'.\exp(-ikx) \quad , \quad x > +l \quad (\text{Eqn A8})$$

Multiplying equations A3 to A6 by A (for $x < -l$) or A' (for $x > +l$) to account for non-unity wave amplitudes on the left hand side of the barrier and comparing with equations A7 and A8 yields the relation:-

$$\begin{aligned} C\exp(ikx) + C'\exp(-ikx) &= A(T_{11}\exp(ikx) + T_{21}\exp(-ikx)) \\ &\quad + A'(T_{12}\exp(ikx) + T_{22}\exp(-ikx)) \end{aligned} \quad (\text{Eqn A9})$$

This relation between the coefficients of the forward and backward propagation terms can be expressed more compactly in matrix form as:-

$$\begin{pmatrix} C \\ C' \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} A \\ A' \end{pmatrix} \quad (\text{Eqn A10})$$

which forms the definition of the transmission matrix. Its function is to relate the coefficients of forward and backward-propagating plane waves to the right of the region to those on the left of the region.

Properties of the transmission matrix.

Under steady-state conditions, the current through the system must be continuous so that no persistent increase in charge occurs. The currents to the left of the barrier must therefore balance those to the right. This requirement results in the identity:-

$$V(k_r) (|C|^2 - |C'|^2) = V(k_l) (|A|^2 - |A'|^2) \quad (\text{Eqn A11})$$

where $V(k_l)$ and $V(k_r)$ are the velocities to the left and the right of the barrier respectively. Equations (A3) and (A4) also show that:-

$$V(\psi_1) = - V(\psi_2) \quad (x < -l) \quad (\text{Eqn A12})$$

so that a forward-propagating plane wave is the complex conjugate of a backward-propagating wave. Applying this result to equations A4 and A6 for $x > l$ gives:-

$$T_{11}^* \exp(-ikx) + T_{21}^* \exp(+ikx) = T_{12} \exp(ikx) + T_{22} \exp(-ikx) \quad (\text{Eqn A13})$$

which shows that

$$T_{11}^* = T_{22} \quad (\text{Eqn A14})$$

$$\text{and } T_{21}^* = T_{12} \quad (\text{Eqn A15})$$

By the use of the T-matrix and the above relations, it is found that the quantity $|C|^2 - |C'|^2$, appearing in the current continuity condition, can be expressed as:-

$$|C|^2 - |C'|^2 = \text{Det}(T) (|A|^2 - |A'|^2) \quad (\text{Eqn A16})$$

and with the help of the current continuity condition (A11), the determinant of the T-matrix is found to be:-

$$\text{Det}(T) = V(k_l)/V(k_r) \quad (\text{Eqn A17})$$

so that for an unbiased barrier, the determinant of the transmission matrix is unity. These results are needed in the calculation of the transmission coefficient of a rectangular barrier.

Relation to the scattering matrix.

An alternative method of describing the interaction of the wave-train with the potential barrier is in terms of a scattering matrix (S-matrix). This relates the coefficients of the plane waves coming out of the region to those going in. Following the same notation as before, the S-matrix is defined as follows:-

$$\begin{pmatrix} C \\ A' \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ C' \end{pmatrix} \quad (\text{Eqn A18})$$

The relation between the S-matrix and the T-matrix can be revealed by solving for C and A' in terms of A and C' from the T-matrix equation with no bias across the barrier.

$$S = (1/T_{22}) \begin{pmatrix} 1 & T_{12} \\ -T_{21} & 1 \end{pmatrix} \quad (\text{Eqn A19})$$

More generally, one should consider the transmitted currents and not just the coefficients and for a biased barrier the above result must be modified to:-

$$S = (1/T_{22}) \begin{pmatrix} f & T_{12} \\ -T_{21} & f \end{pmatrix} \quad (\text{Eqn A20})$$

where f is the ratio of the outgoing to incoming velocities $V(k_r)/V(k_l)$. By forming the product SS^+ , it can be verified that the S-matrix for the currents is unitary, as it must be in order to conserve current. Indeed, one could have started from this condition on the S-matrix to find the appropriate T-matrix in terms of elements of the S-matrix [Borland 1961].

Calculation of the matrix elements for a rectangular potential barrier

Allowing for any difference in effective mass within the barrier and a different wavevector on the right hand side (figure A2), the general form of the wavefunction in each region can be written down.

$$\psi_{1k} = A \exp(ik_1 x) + A' \exp(-ik_1 x) \quad (\text{Eqn A21})$$

$$\psi_{2k} = B \exp(\rho x) + B' \exp(-\rho x) \quad (\text{Eqn A22})$$

$$\psi_{3k} = C \exp(ik_3 x) + C' \exp(-ik_3 x) \quad (\text{Eqn A23})$$

where ρ is the propagation constant inside the barrier. In this calculation it is assumed that the energy of the incident wavetrain is less than the barrier height and hence:-

$$\rho = (2m^* / \hbar^2) (V_{\text{barrier}} - E)^{1/2} \quad (\text{Eqn A24})$$

The corresponding results when the energy is greater than the barrier height can be obtained by letting $\rho \rightarrow i\rho$.

To calculate the matrix elements, consider the following situation in which plane waves are incident on the barrier region from a point far to the left. For the purposes of calculating the matrix elements for a single barrier it is assumed that there are no backward propagating waves on the right hand side. Thus the situation can be thought of one in which the barrier region exists in isolation inside a domain of infinite extent so that an idealised experiment can be performed to find the T-matrix elements. This does not mean that any results so obtained are only valid in the case $C'=0$, since the presence of any backward-propagating waves will not alter the intrinsic transmission properties of the barrier.

In accordance with the previous discussion, the coefficient C' , representing the backward-propagating wave to the right of the barrier is set equal to zero and the forward-propagating wave to the right, C , is set equal to unity. This may not seem like an obvious choice, with the condition $A=1$ seeming more natural, but it will be shown that the choice $C=1$ leads to a particularly simple scheme for finding the matrix elements. Inserting these conditions into the T-matrix equation of (A10) results in:-

$$T_{21} A = -T_{22} A' \quad (\text{Eqn A25})$$

$$T_{11} A + T_{12} A' = 1 \quad (\text{Eqn A26})$$

Substituting A25 into A26 gives

$$A (T_{22} T_{11} - T_{12} T_{21}) = T_{22} \quad (\text{Eqn A27})$$

For no bias, it was shown that the determinant of the T-matrix was unity. Therefore,

$$T_{22} = A \quad (\text{Eqn A28})$$

This shows that if the coefficient A can be found from equations A21 to A23 then the matrix element T_{22} has also been found. Similarly, it is also found that:-

$$A' = -T_{21} \quad (\text{Eqn A29})$$

and the other elements can be found using the property expressed by relations A14 and A15. The transmission coefficient for the barrier is defined as the ratio of the transmitted to incident fluxes.

$$T = |C|^2 / |A|^2 \quad (\text{Eqn A30})$$

so that having normalised the transmitted wave amplitude to unity, the transmission coefficient can be written as:-

$$T = 1 / |T_{22}|^2 \quad (\text{Eqn A31})$$

For a barrier under bias, one must ensure that the current is normalised by taking the outgoing C component to be $(V(k_r)/V(k_l))^{1/2}$ and writing the determinant of the T-matrix in its more general form (equation A17). These extensions to the unbiased transmission result in a modified transmission formula:-

$$T(k) = (k_r/k_l) 1/|T_{22}|^2. \quad (\text{Eqn A32})$$

The method for finding the transmission coefficient of a rectangular potential barrier therefore, is to solve for the coefficient A (and hence T22) between equations A21 to A23. This is achieved using the continuity conditions for the wavefunction at each interface, namely that the probability density and the current must be continuous at each discontinuity. By matching the envelope functions, $\psi(x)$ and $1/m^* d\psi/dx$ across each interface the resulting set of equations for $C=1$ and $C'=0$ can be solved to yield:-

$$B = 1/2 ((1 + ik_3)/(\rho\lambda)) \exp((ik_3 - \rho)l) \quad (\text{Eqn A33})$$

$$B' = 1/2 ((1 - ik_3)/(\rho\lambda)) \exp((ik_3 + \rho)l) \quad (\text{Eqn A34})$$

$$A = 1/2((k_1 + k_3)/k_1) \cosh(2\rho l) - i((k_1 k_3 - \lambda^2 \rho^2)/(\rho\lambda k_1)) \sinh(2\rho l) \\ * \exp(i(k_3 + k_1)l) \quad (\text{Eqn A35})$$

$$A' = 1/2((k_1 - k_3)/k_1) \cosh(2\rho l) - i((k_1 k_3 + \lambda^2 \rho^2)/(\rho\lambda k_1)) \sinh(2\rho l) \\ * \exp(i(k_3 - k_1)l) \quad (\text{Eqn A36})$$

where the difference in effective mass in the barrier and well regions is characterised by the parameter λ .

$$\lambda = m^*(\text{well}) / m^*(\text{barrier}) \quad (\text{Eqn A37})$$

Figure A1

Type of potential considered in transmission-matrix calculations.

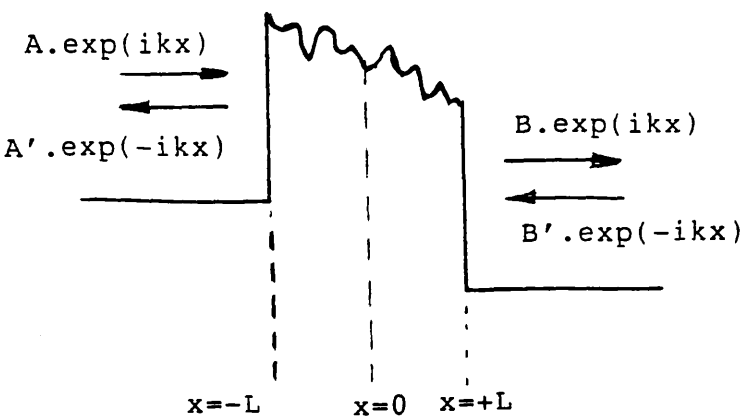
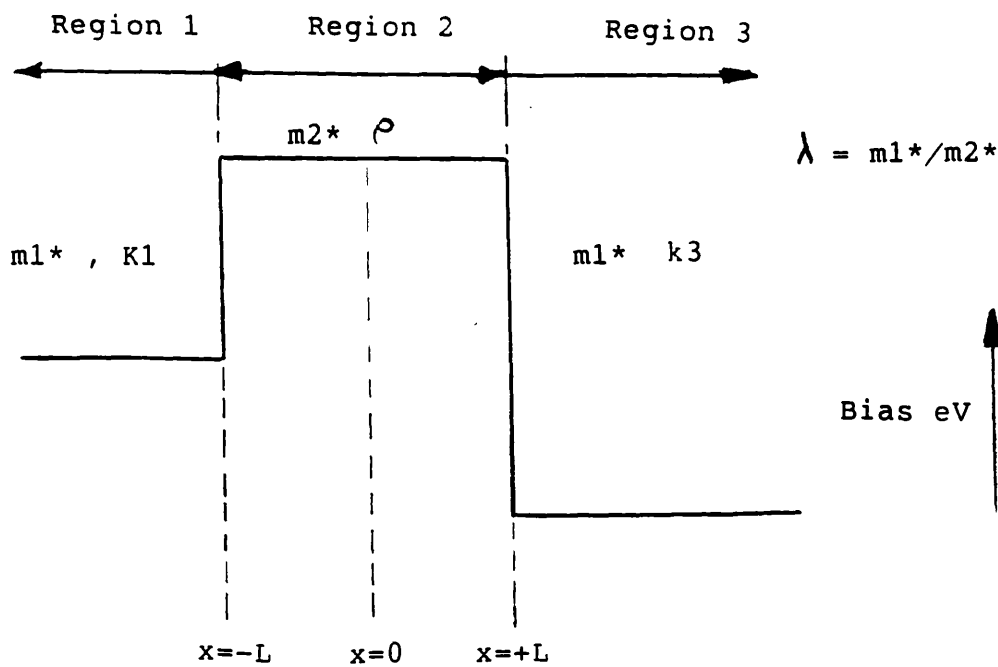


Figure A2 Model potential used to calculate T-matrix elements analytically.



Appendix B: The stationary phase approximation.

This appendix outlines the stationary phase approximation used to calculate tunnelling delay times for the resonant double barrier discussed in chapter 2.

The stationary phase approximation can be derived by writing the complex transmission coefficient as a modulus and phase.

$$t(k) = |t(k)| \exp(i s(k)) \quad (\text{Eqn B1})$$

Ultimately, this complex transmission coefficient can be thought of as completely characterising the passage of the electron through the structure. When the wavepacket has fully interacted with the barrier, the transmitted wavefunction can be constructed from the Fourier components, $g(k)$, of the initial wavepacket by multiplying each component by the complex transmission coefficient.

$$\psi(x,t) = \int g(k) t(k) \exp(i(-kx - \omega t)) dk \quad (\text{Eqn B2})$$

The stationary phase approximation then consists in Taylor-expanding the phase of the transmission, $s(k)$, about some central wavevector k_0 to terms linear in k .

$$s(k) = s(k_0) + (k - k_0) ds/dk|_{k=k_0} \quad (\text{Eqn B3})$$

Substituting this into equation B2 yields:-

$$\psi(x,t) = \exp(i(s(k_0) + k_0 x)) \int |t(k)| \exp(-i(k - k_0)(x + ds/dk) - i\omega t) dk \quad (\text{Eqn B4})$$

This can now be compared with the standard form of a wavepacket, centred on x_0 in real space and k_0 in k -space. It is seen that for any values of x or k too far away from the central values, the exponential starts to oscillate so rapidly that the k -integration averages ψ to zero, so that the transmitted wavefunction can be written approximately as:-

$$\psi(x,t) \propto \int |t(k)| \exp(-i(k - k_0)(x - x_0) - i\omega t) dk \quad (\text{Eqn B5})$$

The new centre of the wavepacket has therefore been shifted due to its passage through the barrier system, and x_0 can be identified with $-ds/dk$. This means that the centre has been set backwards by an amount ds/sk . Assuming that the average (group) velocity of the wavepacket has not otherwise been changed, the additional time required to travel through the barrier region is:-

$$\tau_d = (1/V(k_0)) \, ds/dk \quad (\text{Eqn B6})$$

and the total time is:-

$$\tau_d = (1/V(k_0)) \, (w + ds/dk) \quad (\text{Eqn B7})$$

where w is the barrier width and $V(k_0)$ is the group velocity of the free wavepacket centred on wavevector k_0 . Assuming a parabolic energy-momentum relationship, the delay time is given by:-

$$\tau_d = \hbar \, ds/dE \quad (\text{Eqn B8})$$

so that in this approximation the delay time is directly proportional to the energy gradient of the phase of the complex transmission coefficient. This is the formula used to calculate the tunnel delay time from the complex transmission coefficient in chapter 2.

Appendix C: Estimation of wavepacket energy from the sheet carrier concentration.

This appendix describes how the energy of the wavepacket used in the computer simulations is estimated from the quoted two-dimensional electron concentration. It is also shown that the 2D density of states function could be significantly changed in thin conductors. For the AB ring modelled in this thesis however, this is not a significant effect.

The total number of electrons in the 2DEG covering an area L^2 is given by:-

$$N_{tot} = (\pi K_f^2) g_s / (2\pi/L)^2 \quad (\text{Eqn C1})$$

where g_s , the spin degeneracy factor, is equal to 2 in the absence of a magnetic field. The number of electrons per unit area (N_{tot}/L^2), or sheet concentration is thus:-

$$n_s = K_f^2 / 2\pi \quad (\text{Eqn C2})$$

and the density of states function, $g(E)$, is therefore given by:-

$$(1/2\pi) 2m^*/\hbar^2 \quad (\text{Eqn C3})$$

If it is assumed that conduction occurs via electrons close to the Fermi energy, then the appropriate energy for the wavepacket is:-

$$E = \pi\hbar^2 n_s / m^* \quad (\text{Eqn C4})$$

Typical sheet concentrations for GaAs/AlGaAs heterojunctions vary from about 2.0×10^{15} to $5.0 \times 10^{15} \text{ m}^{-2}$ corresponding to Fermi energies of between 7 and 18 meV. These energies are measured with respect to the bottom of the two-dimensional conduction band. The energy with respect to the three-dimensional conduction band for the system as a whole is 7 to 18 meV the electrons have in the two-dimensional conduction band, plus the energy due to the confinement of the wavefunction perpendicular to the interface.

The energy levels due to confinement perpendicular to the interface may be approximately calculated by assuming that the potential profile of the well is triangular with an infinitely high potential at the heterojunction interface. Application of Gauss' theorem to a thin sheet of charge located just to the right of the interface results in an electric field strength of:-

$$F_z = e n_s/(\epsilon_r \epsilon_0) \tag{Eqn C5}$$

so that the Schrodinger equation for electrons moving in this potential (in one dimension) is:-

$$(-\hbar^2/2m^*) \, d^2\psi/dz^2 - eF_z z \psi = E_n \, \psi \tag{Eqn C6}$$

This gives a solution for the energy levels of [Stern 1972, Abromowitz and Stegun 1984]

$$E_n = (\hbar^2/2m^*)^{1/3} (eF_z)^{2/3} a_n \tag{Eqn C7}$$

where the a_n are the zeros of the Airy function.

Taking the relative permittivity of GaAs to be 12.9 and an effective mass of 0.067 times the free electron mass, gives the following values for the first two sub-bands.

$$n_s = \qquad 2.0 \times 10^{15} \, m^{-2} \qquad 5.0 \times 10^{15} \, m^{-2}$$

$$E_1 = \qquad 38.6 \, meV \qquad 71.0 \, meV$$

$$E_2 = \qquad 67.4 \, meV \qquad 124.1 \, meV$$

The energy separation between the first and second sub-bands is therefore large compared to the typical energies of about 10meV within the two-dimensional plane, so that only one sub-band need be considered at low temperatures.

In addition to the confinement in the z-direction (perpendicular to the interface) there can also be significant confinement effects in the two-dimensional plane if the conducting channels are thin enough. If the width of the conducting channels within the

2DEG become commensurate with the DeBroglie wavelength then the energy levels due to the transverse confinement may become significantly large.

The following calculation sets out to determine what effect this additional confinement has on the Fermi-energy compared to that computed from a purely two-dimensional density-of-states function. The density-of-states function for a purely one-dimensional channel is:-

$$g(E) = (g_s/(2\pi)) (2m^*/\hbar^2)^{1/2} E^{-1/2} \tag{Eqn C8}$$

Although this formula applies to strictly one-dimensional conduction, the conducting channel can be of any width provided that conduction is restricted to only a single 1D sub-band. However, in wires of realistic width, there will be a number of occupied 1D sub-bands, and so one must consider the sum of the all the 1D density-of-states over all sub-bands.

$$g(E) = \sum (g_s/2\pi) (2m^*/\hbar^2)^{1/2} (E-E_s)^{-1/2} H(E-E_s) \tag{Eqn C9}$$

Where s is the sub-band index and $H(x)$ the Heaviside step function. The ground-state energy for each sub-band, E_s , was calculated from the particle-in-a-box eigenfunctions for an infinitely deep potential well. Starting from the density of states function for the quasi-1D system, one can always define a Fermi energy from the following relation for the number density [Blakemore 1985]:-

$$n = \int_0^\infty g(E) \{ 1 + \exp[(E-E_f)/(K_bT)] \}^{-1} dE \tag{Eqn C10}$$

In this work however, one is usually only interested in very low temperatures of the order of a few Kelvin at most, so the above form may be simplified to:-

$$n = \int_0^{E_f} g(E) dE \tag{Eqn C11}$$

Figures C1 to C3 show the calculated density of states for conductor widths of 20nm, 66.4nm and 200nm respectively after an adjustment of units to an areal density had been made for the appropriate wire width.

The density of states for the 20nm wire has a very one-dimensional character with only one sub-band occupied up to an energy of 56meV. As the wire width increased the density of states contained contributions from a greater number of 1D sub-bands and for conductor widths of 200nm, the density of states can be seen to be converging to the two-dimensional result.

However, in this thesis, it was only how this deviation from the purely two-dimensional behaviour affected the relation between carrier density and Fermi energy which was of major concern. The integral in equation C1 was therefore evaluated for several values of Fermi-energy using the appropriate quasi-1D density of states function.

Integration over energy had the effect of smoothing-out the effects of the sharp peaks in the density of states. For small conductor widths of about 20nm, there is a considerable discrepancy between the result using the quasi-1D density of states and that using the purely two-dimensional density of states (figure C4).

Figures C5 and C6 show the results of the same calculation for conductor widths of 66.4nm and 200nm, 66.4nm being the width used in the initial full-size computer simulation of an AB ring. Although the structure from the quasi-1D density of states can be seen in the curve for the 66.4nm conductor, the actual values of carrier density as a function of the Fermi energy did not differ greatly from the 2D result, the error in the Fermi energy at a carrier density of $5 \times 10^{15} \text{ m}^{-2}$ being about 12%.

The conclusion to be drawn from these results is that for conductor widths of current interest (80-200nm), the quasi-1D nature of the density of states does not have a large effect on the calculated Fermi energy. Set in the context of experimental uncertainties regarding the carrier concentration, conductor width and the assumptions made in constructing the computer model, these corrections can probably be neglected, although it is useful to be aware of their existence particularly if very thin conductors are to be considered.

Figure C1

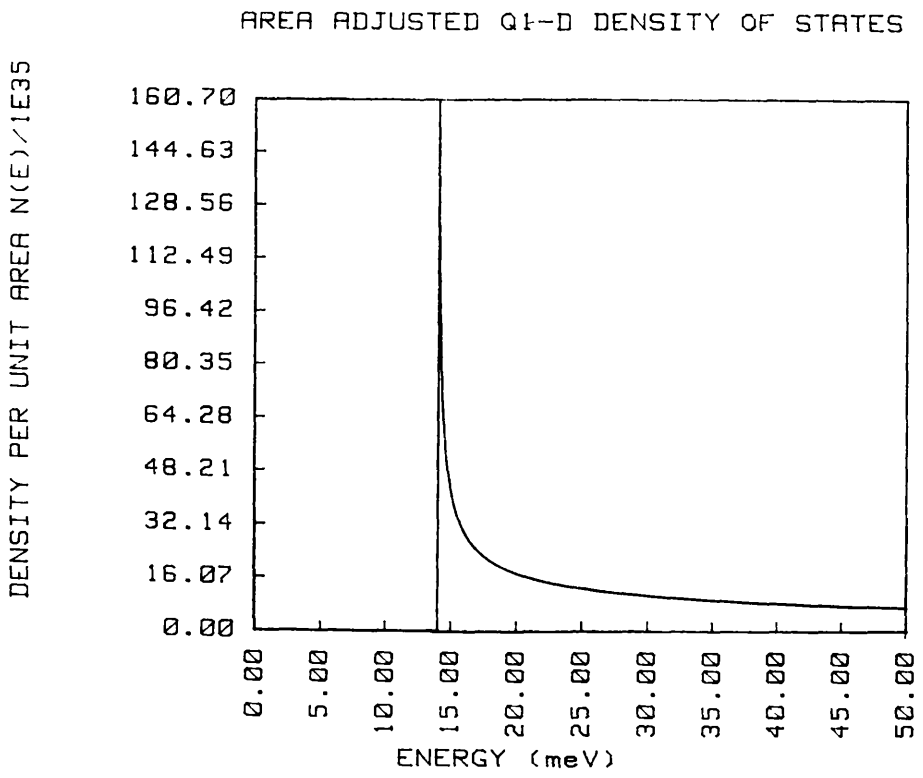


Figure C2

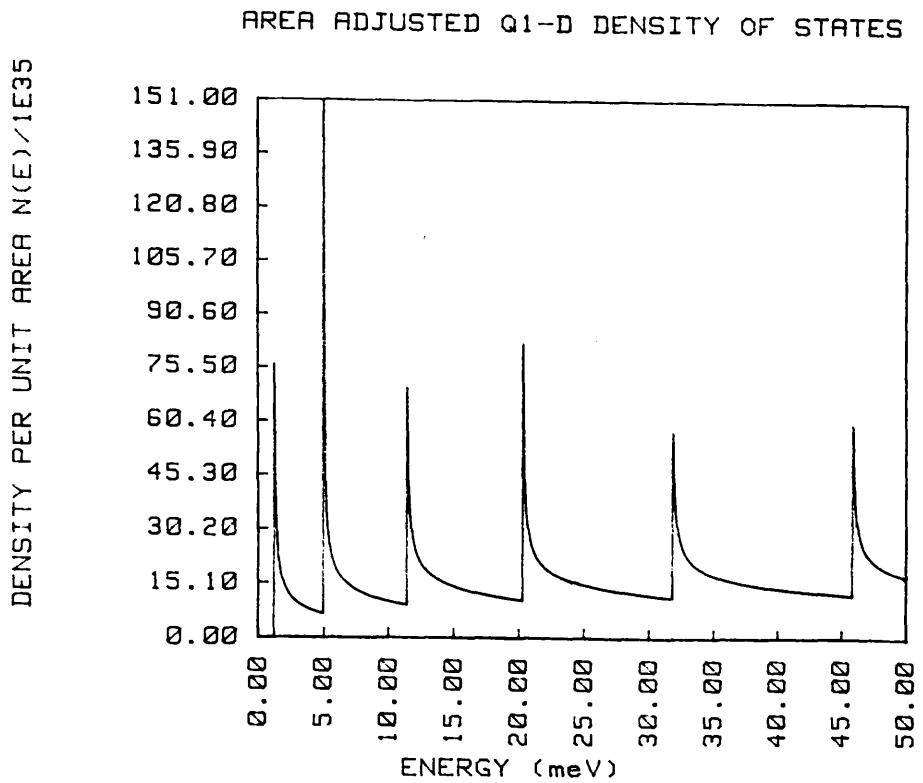


Figure C3

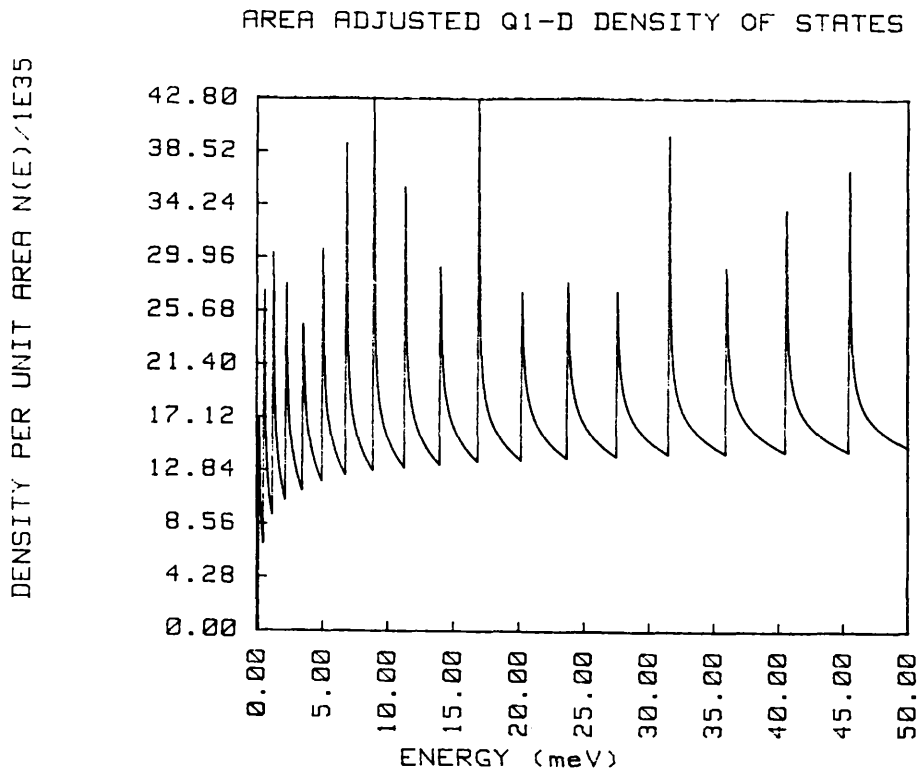


Figure C4

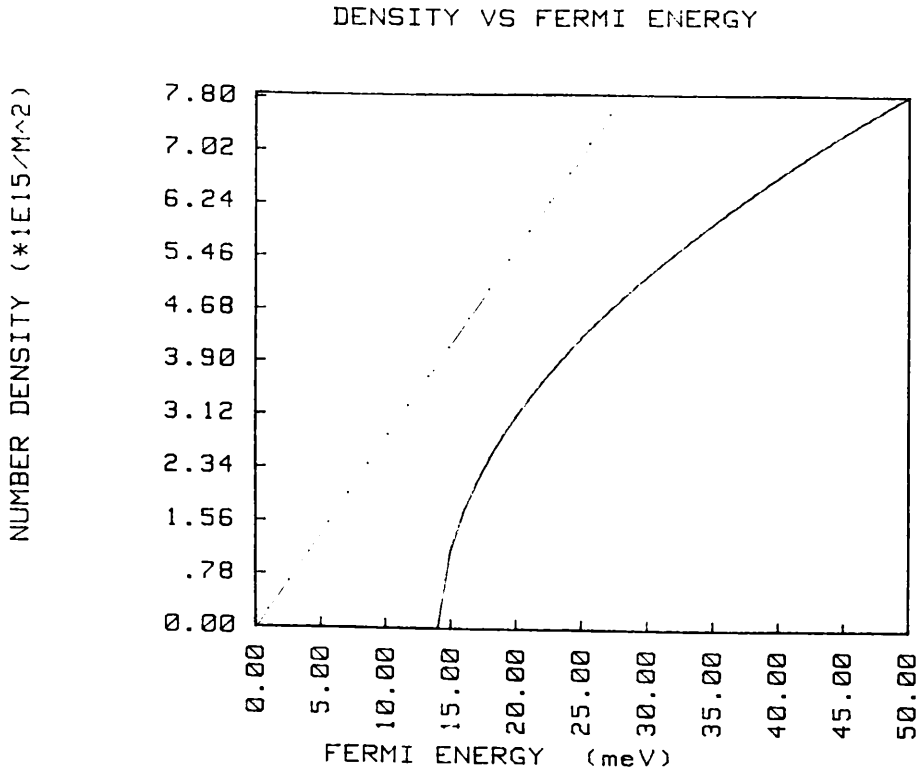


Figure C5

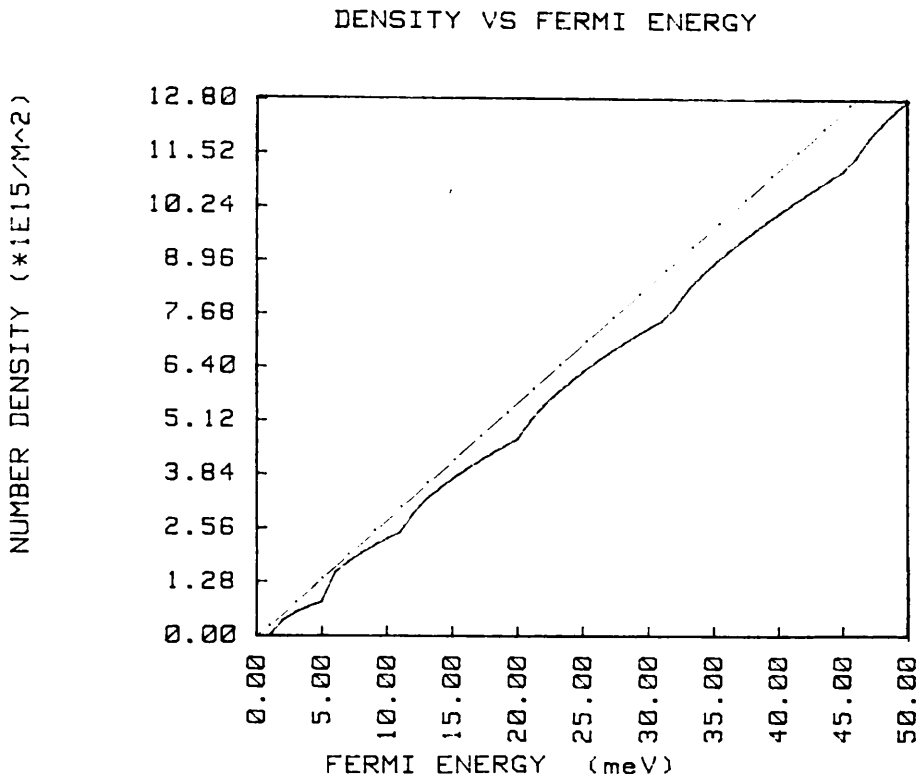
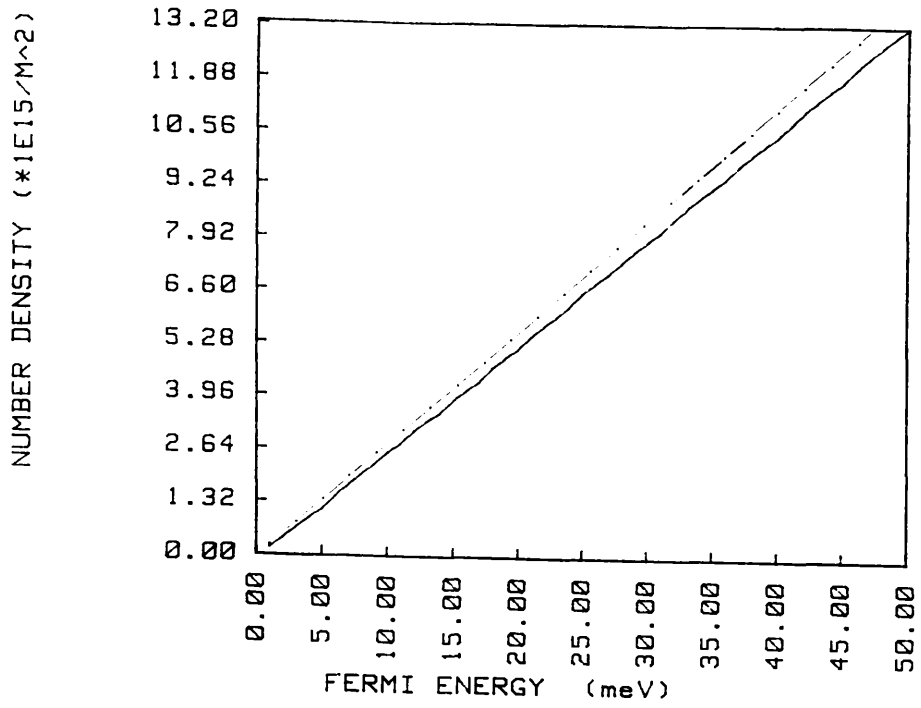


Figure C6

DENSITY VS FERMI ENERGY



References.

- Abromowitz M and Stegun I.A (1984), Pocketbook of mathematical functions (Verlag Harri Deutsch).
- Aharonov Y and Bohm D (1959), Phys. Rev. Lett. **115** 485.
- Al-Mudares M (1984), Ph.D Thesis, University of Surrey.
- Al-Mudares M (1987), Private communication.
- Al'tshuler B.L, Aronov A.G and Spivak B.Z (1981), JETP Lett. **33** 94.
- Bandyopadhyay S, Melloch M.R, Datta S, Das B, Cooper J.A and Lundstrom M.S (1986), Proceedings of the International Electronic Devices Meeting, Los Angeles CA. Dec 7-10 p76.
- Barker J.R and Murray S (1983), Phys. Lett. **93A** 271.
- Barker J.R (1985), Physica **134B** 22.
- Barker J.R (1986) in Physics and Fabrication of microstructures and microdevices, Ed M.J Kelly and C. Weisbuch (Springer, Berlin 1986).
- Barker J.R and Laughton M (1989), To be published.
- Berggren K.F Thornton T.J Newson D.J and Pepper M (1986), Phys. Rev. Lett. **57** 1769.
- Berry M.V (1984), Proc. Roy. Soc. London A **392** 45.
- Berry M.V (1987) at the Institute of Physics Conference, Bristol 1987.
- Blakemore J.S (1985), Solid state Physics 2nd Ed. (Cambridge University Press).
- Borland R.E (1961), Proc. Phys. Soc. **78** 926.
- Broers A.N (1986) in Physics and Fabrication of microstructures and microdevices, Ed M.J Kelly and C. Weisbuch (Springer, Berlin 1986).
- Buttiker M, Imry Y and Landauer R (1983), Physics Lett. **96A** 365.
- Buttiker M, Imry Y and Azbel M. Ya (1984), Phys. Rev. A **30** 1982.
- Buttiker M, Imry Y, Landauer R and Pinhas S (1985), Phys. Rev. B **31** 6207.
- Buttiker M (1986) in proceedings of the conference on new techniques and ideas in quantum measurement theory, New York 1986.
- Cahay M, McLennan M and Datta S (1988), Phys. Rev. B **37** 10125.
- Capasso F, Mohammed K and Cho A. Y (1986), IEEE J. Quantum electronics **QE22** 1853.
- Chambers R.G (1960), Phys. Rev. Lett. **5** 3.
- Chiao R.Y and Wu Y.S (1986), Phys. Rev. Lett. **57** 933.
- Colella R.A, Overhauser A.W and Werner S.A (1975), Phys. Rev. Lett **34** 1472.
- Collins S 1986, Ph.D Thesis University of Warwick.
- Datta S and Bandyopadhyay S (1987), Phys. Rev. Lett. **58** 717.

- Davies J.H (1988), *Semicond. Sci. Technol.* **3** 995.
- Economou and Soukoulis (1980), *Phys. Rev. B* **22** 3519.
- Ehrenberg W and Siday R.E (1949), *Proc. Phys. Soc. London, Sect B* **62** 8.
- Esaki L (1986), *IEEE J. Quantum electronics* **QE22** 1611.
- Finch M (1987), Institute of Physics Conference, Bristol (unpublished).
- Fisher and Lee (1981), *Phys. Rev. B* **23** 6851.
- Ford C.J.B, Thornton T.J, Newbury R, Pepper M, Ahmed H, Davies G.J and Andrews D (1987), *Proceedings from the 3rd International Conference on superlattices , microstructures and microdevices*, Chicago IL 17-20 August 1987.
- Ford C.J.B, Thornton T.J, Newbury R, Pepper M, Ahmed H, Foxton C.T, Harris J.J and Roberts C (1988), *J. Phys. C Solid State Phys.* **21** L325.
- Ford C.J.B (1988), Private communication.
- Ford C.J.B, Thornton T.J, Newbury R, Pepper M, Ahmed H, Peacock D.C, Ritchie D.A, Frost J.E.F and Jones G.A.C (1989), *Appl. Phys. Lett.* **54** 21.
- Franz W (1939), *Verh. Dtsch. Phys. Ges* (2) 65.
- Fresbach H and Villars F (1958), *Rev. Mod. Phys.* **30** 24.
- Gefen Y, Imry Y and Azbel M. Ya (1984), *Phys. Rev. Lett.* **52** 129.
- Gijs M, Van Haesendonck C and Bruynseraede Y (1984), *Phys. Rev. Lett.* **52** 2069.
- Goldberg A, Schey H.M and Schwartz J.L (1967), *Am. J. Phys.* **35** 177.
- Goldstein H (1959), *Classical mechanics* (Reading Mass).
- Hartstein A (1986) in *Physics and Fabrication of microstructures and microdevices*, Ed M.J Kelly and C. Weisbuch (Springer, Berlin 1986).
- Ishibashi K, Takagaki Y, Gamo K, Namba S, Ishida S, Murase K, Aoyagi Y and Kawabe M (1987), *International Symposium on Anderson localisation*, University of Tokyo 1987.
- Iwata N, Matsumoto Y, Baba T and Ogawa M (1986), *Jap. J. Appl. Phys.* **25** L249.
- Jaros M (1986) in *Physics and Fabrication of microstructures and microdevices*, Ed M.J Kelly and C. Weisbuch (Springer, Berlin 1986).
- Kobe D.H and Yang K (1988), *Am. J. Phys.* **56** 549.
- Landau L.D and Lifshitz E.M, *Quantum mechanics*, Pergamon Press (1958).
- Landauer R (1957), *IBM J. Res. Dev.* **1** 223.
- Laux S.E, Frank D.J and Stern F (1988), *Surf. Sci.* **196** 101.
- Leadbeater M.L, Taylor R.P, Main P.C, Eaves L, Beaumont S.P, McIntyre I, Thoms S and Wilkinson C.D.W (1987), *Proceedings from the 14th International Conference on GaAs and related compounds*, Crete 1987.
- Lee P.A and Fisher D.S (1981), *Phys. Rev. Lett* **47** 882.
- Lee P.A and Stone A.D (1985), *Phys. Rev. Lett* **55** 1622.

Lenstra D and Van Haeringen W (1986), Phys. Rev. Lett. **57** 1623.
 Mandelung E (1926), Z. Phys. **40** 322.
 McIntyre I (1988), Private Communication, University of Glasgow.
 Messiah A (1962), Quantum mechanics, North-Holland.
 Mollenstedt G and Duker H (1956), Z. Phys, **145** 377.
 Olariu S and I.I Popescu (1985), Rev. Mod. Phys. **57** 339.
 Pepin J (1987), Private communication.
 Pepin J (1989), To be published.
 Peres A (1983), J. Math. Phys. **24** 1110.
 Potter D (1973), Computational Physics (Wiley and Sons).
 Press W.H, Flannery B.P, Teukolsky S.A and Vetterling W.T (1987), Numerical recipes (Cambridge University Press).
 Razeghi M, Tordella A, Davies R.A, Long A.P, Kelly M.J, Britton E, Boothroyd C and Stobbs W.M (1987), Electronics Lett. **23** 116.
 Schiff L.I (1955), Quantum mechanics 2nd Ed.
 Sharvin D. Yu and Sharvin Yu.V (1981), JETP Lett. **34** 273.
 Skocpol W.J (1986) in Physics and Fabrication of microstructures and microdevices, Ed M.J Kelly and C. Weisbuch (Springer, Berlin 1986).
 Sollner T.C.L.G, Goodhue W.D, Tannenwald P.E, Parker C.D and Peck D.D (1983), Appl. Phys. Lett. **43** 588.
 Stern F (1972), Phys. Rev. B **5** 4891.
 Stevens K (1983) J.Phys. C **16** 3649.
 Stone A.D (1985), Phys. Rev. Lett **54** 2692.
 Stone A.D and Imry Y (1986), Phys. Rev. Lett **56** 375.
 Stone A.D and Szafer A (1988), IBM J. Res. Dev. **32** 384.
 Taylor R.P (1988) at the University of Glasgow, Unpublished.
 Thornton T.J, Pepper M, Ahmed H, Andrews D and Davies G.J (1986), Phys. Rev. Lett. **56** 1198.
 Thouless D.J (1980), Solid State Comm. **34** 683.
 Timp. G, Chang A.M, Cunningham J.E, Chang T.Y, Mankiewich P, Behringer R and Howard R.E (1987), Phys. Rev. Lett **58** 2814.
 Tomita A and Chiao R.Y (1986), Phys. Rev. Lett **57** 937.
 Tsu R and Esaki L (1973), Appl. Phys. Lett. **22** 562.
 Umbach C.P, Washburn S, Laibowitz R.B and Webb R.A (1984), Phys. Rev. B **30** 4048.
 Van Houten H, Van Wees B.J, Heijmann M.G.J and Andre J.P (1986), Appl. Phys. Lett **49** 1781.

- Vigneron J.P and Lambin P.H (1980), J. Phys. A **13** 1135.
- Washburn S and Webb R.A (1986), Advances in Physics **35** 375.
- Webb R.A, Washburn S, Umbach C.P and Liabowitz R.B (1985), Phys. Rev. Lett. **54** 2696.
- Wharam D.A, Thornton T.J, Newbury R, Pepper M, Ahmed H, Frost J.E.F, Hasko D.G, Peacock D.C, Ritchie D.A and Jones G.A.C (1988), J.Phys.C **21** L209.
- Wind S, Rooks M.J, Chandrasekhar V and Prober D.E (1986), Phys. Rev. Lett. **57** 633.

